



Caractérisation objective et modélisation psychovisuelle du flou de mouvement sur les écrans à cristaux liquides - Impact sur la qualité perçue

Sylvain Tourancheau

► To cite this version:

Sylvain Tourancheau. Caractérisation objective et modélisation psychovisuelle du flou de mouvement sur les écrans à cristaux liquides - Impact sur la qualité perçue. Traitement du signal et de l'image [eess.SP]. Université de Nantes, 2009. Français. NNT : . tel-00458786

HAL Id: tel-00458786

<https://theses.hal.science/tel-00458786>

Submitted on 22 Feb 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE NANTES

ÉCOLE DOCTORALE

« SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE MATHÉMATIQUES »

Année : 2009

Thèse de Doctorat de l'Université de Nantes

Spécialité : AUTOMATIQUE, ROBOTIQUE, TRAITEMENT DU SIGNAL
ET INFORMATIQUE APPLIQUÉE

présentée et soutenue publiquement par

Sylvain TOURANCHEAU

le 3 décembre 2009

à l'École polytechnique de l'université de Nantes

**Caractérisation objective et modélisation psychovisuelle
du flou de mouvement sur les écrans à cristaux liquides.
Impact sur la qualité perçue**

Jury

Rapporteurs	M. Jean-Marc Chassery	<i>Directeur de Recherche CNRS, GIPSA-lab, Grenoble</i>
	M. Azeddine Beghdadi	<i>Professeur, L2TI, Université Paris 13</i>
Examineurs	Mme Ingrid Heynderickx	<i>Full Professor, Delft University of Technology, Pays-Bas</i>
	M. Kjell Brunnström	<i>Senior Scientist, Acreo AB, Kista, Suède</i>
	M. Dominique Barba	<i>Professeur émérite, IRCCyN, Polytech'Nantes</i>
	M. Patrick Le Callet	<i>Professeur, IRCCyN, Polytech'Nantes</i>

Directeur de Thèse : Dominique Barba

Co-encadrant : Patrick Le Callet

Laboratoire : Institut de Recherche en Communications et en Cybernétique de Nantes (IRCCyN)

Composante de rattachement du directeur de thèse : École polytechnique de l'université de Nantes

The revolution will not be televised.

GIL SCOTT-HERON

Remerciements

Je tiens à remercier tout particulièrement Patrick Le Callet, qui par son soutien permanent m’a permis de mener à bien ces travaux de thèse. J’ai apprécié la façon avec laquelle il a encadré mon travail, veillant toujours à ce que je ne perde pas de vue l’objectif final, tout en me laissant une autonomie et une liberté suffisantes pour exprimer mes propres idées. J’ai pu compter sur lui à tout moment, que ce soit pour me remotiver dans les périodes de doute, ou pour me conseiller lorsqu’il a fallu prendre certaines décisions. Au delà d’un encadrement de thèse, il a su me former au métier de chercheur. C’est un héritage dont je suis fier et que j’ai hâte de pouvoir transmettre à mon tour.

Je remercie également Dominique Barba, qui malgré une retraite bien méritée a su se montrer disponible et efficace. J’ai notamment apprécié la rapidité et la rigueur de ses relectures durant les semaines qui ont précédé le dépôt du manuscrit.

Je remercie chaleureusement Jean-Marc Chassery et Azeddine Beghdadi d’avoir accepté d’être rapporteurs de cette thèse. J’ai apprécié leurs remarques pertinentes et leurs commentaires constructifs. Je remercie également Ingrid Heynderickx de m’avoir fait l’honneur de présider ce jury. Enfin, je tiens à remercier Kjell Brunnström d’avoir accepté d’être examinateur de ma thèse, mais surtout d’avoir été l’instigateur puis le moteur d’une collaboration qui dure depuis maintenant presque 3 ans. Les travaux présentés dans cet ouvrage lui doivent beaucoup. *Tack så väldigt mycket!*

Je remercie également ceux sans qui cette thèse ne serait pas ce qu’elle est : Arnaud Tirel et Mathieu Carnec tout d’abord, engagés avec moi sur le projet européen HD4U, puis Romuald Pépion et Romain Cousseau qui leur ont succédé. Leur collaboration au niveau du développement logiciel et de l’organisation des tests subjectifs fut précieuse.

Je remercie les doctorants qui m’ont accompagné durant cette thèse, leur soutien et leurs encouragements illustrent parfaitement la légendaire solidarité entre thésards. Je tiens à saluer plus particulièrement ceux de ces compagnons de route qui ont su me supporter quotidiennement pendant ces quelques années : Olivier, Eddy, Guillaume, Stéphane, Fadi et Éloïse.

Je remercie tous les autres membres de l’équipe Image Video Communications du laboratoire IRC-CyN. Ces quelques années avec vous ont été plus qu’agréables et l’ambiance chaleureuse et fraternelle qui règne au sein de l’équipe est également un facteur de réussite important.

Enfin, je remercie ma famille, mes amis, et surtout Clarisse, pour tout le reste...

Table des matières

Introduction générale	1
 I Visualisation d'images animées, du CRT au LCD, concepts technologiques et nouveaux enjeux	 7
1 Systèmes d'affichage électronique, évolution et limites du CRT	9
1.1 Introduction	9
1.2 Concepts de base	10
1.2.1 Structure d'un système d'affichage	10
1.2.2 Notions et terminologie	11
1.2.3 Distinction entre écran et contenu	13
1.3 Télévision et écran à tube cathodique	13
1.3.1 Naissance de la télévision	13
1.3.2 Écran à tube cathodique (CRT)	14
1.3.3 Standardisation de la télévision analogique	16
1.4 Vers une amélioration de la qualité d'usage	16
1.4.1 Immersion, impact et expérience comparable au cinéma	16
1.4.2 Limites de l'écran à tube cathodique et de la télévision standard	17
1.4.3 Conditions pour une qualité d'usage accrue	18
1.4.4 Nouveaux écrans	19
1.5 Conclusion	21
 2 Écrans à cristaux liquides, un nouveau contexte	 23
2.1 Introduction	23
2.2 Les cristaux liquides	24
2.2.1 La découverte des cristaux liquides	24
2.2.2 Le cristal liquide en quelques mots	24
2.2.3 Application aux écrans	25

2.3	Fonctionnement d'un écran à cristaux liquides	25
2.3.1	Le panneau d'affichage	26
2.3.2	L'éclairage	29
2.3.3	L'adressage des pixels	34
2.4	Une nouvelle technologie à évaluer	35
2.4.1	De nouvelles conditions de visualisation	35
2.4.2	De nouvelles distorsions visuelles	37
2.4.3	Problématique	38
2.5	Conclusion	39

II Mesures photométriques du flou de mouvement sur les écrans à cristaux liquides 41

3 Le flou de mouvement sur les écrans à cristaux liquides 43

3.1	Introduction	43
3.2	Flou de mouvement	44
3.2.1	Causes	44
3.2.2	Formation du flou de mouvement	45
3.2.3	Méthodes de réduction du flou de mouvement	47
3.3	Réponse temporelle indicielle d'un écran LCD	48
3.3.1	Mesure de la réponse indicielle	49
3.3.2	Le temps de réponse	51
3.3.3	Normalisations concernant le temps de réponse des écrans LCD	54
3.4	Profil spatial d'un contour en mouvement	57
3.4.1	Systèmes de mesures du flou de mouvement par dispositifs imageurs	57
3.4.2	Mesure du flou de mouvement	58
3.4.3	Largeur de flou	61
3.4.4	Valeur globale du flou de mouvement pour un écran	63
3.5	Conclusion	64

4 Contribution à la mesure du flou de mouvement sur LCD 65

4.1	Introduction	65
4.2	Analyse de la formation du flou de mouvement sur le capteur	66
4.2.1	Signal affiché	67
4.2.2	Déplacement du capteur	67
4.2.3	Intégration temporelle	67

4.2.4	Conséquence	69
4.3	Mesures du flou de mouvement	69
4.3.1	Écrans testés	69
4.3.2	Mesures	71
4.4	Résultats	73
4.4.1	Observations	75
4.4.2	Comparaison des résultats obtenus par chacune des deux méthodes	76
4.4.3	Relation entre temps de réponse et temps de flou	78
4.5	Conclusion	79
III	Étude psychophysique du flou de mouvement	81
5	Influence du mouvement sur la réponse supraliminaire du système visuel humain	83
5.1	Introduction	83
5.2	Contexte	84
5.2.1	Fonctions de sensibilité au contraste	84
5.2.2	Fonctions de sensibilité au contraste en mouvement	85
5.2.3	Fonctions de sensibilité supraliminaires	86
5.3	Expérimentations psychophysiques	87
5.3.1	Observateurs	87
5.3.2	Matériel	87
5.3.3	Stimuli et conditions de visualisation	87
5.3.4	Procédure	89
5.4	Résultats	90
5.4.1	Données mesurées	90
5.4.2	Modélisation du gain	92
5.4.3	Surface spatio-temporelle du gain en sensibilité	94
5.5	Discussion	95
5.5.1	Stabilisation du contraste	95
5.5.2	Comparaison avec les mesures effectuées au seuil de visibilité	98
5.6	Conclusion	100
6	Perception du flou de mouvement	103
6.1	Introduction	103
6.2	Travaux antérieurs	104
6.3	Description de l'expérience	105

6.3.1	Mesures physiques	105
6.3.2	Protocole	105
6.3.3	Écrans testés	107
6.3.4	Conditions de visualisation	107
6.3.5	Observateurs	108
6.4	Résultats et interprétation	108
6.4.1	Résultats	108
6.4.2	Comparaison avec les autres travaux	109
6.4.3	Étude de la variation du coefficient A	111
6.4.4	Évaluation globale du modèle	115
6.5	Vers une mesure perceptive du flou de mouvement	117
6.6	Conclusion	118
7	Gêne visuelle associée au flou de mouvement	121
7.1	Introduction	121
7.2	Expérimentations	123
7.2.1	Écrans et conditions de visualisation	123
7.2.2	Séquences	123
7.2.3	Procédure	125
7.2.4	Observateurs	125
7.3	Résultats	126
7.3.1	Détermination des vitesses seuil	126
7.3.2	Conversion des vitesses seuil	126
7.4	Observations et discussion	132
7.4.1	Valeurs de flou de mouvement aux vitesses seuil	132
7.4.2	Conséquences sur la valeur recommandée du temps de flou perçu global . . .	135
7.5	Conclusion	137
IV	Impact du flou de mouvement sur la qualité visuelle	139
8	Évaluation subjective de la qualité visuelle sur les écrans LCD et CRT	141
8.1	Introduction	141
8.2	Tests d'évaluation subjective de la qualité	142
8.2.1	Méthodologie	142
8.2.2	Contenus évalués	143
8.2.3	Conditions d'observation	147

8.2.4	Observateurs	148
8.2.5	Résultats	149
8.3	Différence de qualité entre écran CRT et écran LCD	150
8.3.1	Cas des séquences vidéo	153
8.3.2	Cas des images fixes	155
8.3.3	Influence de la résolution	156
8.4	Conclusion	159
9	Vers une prédiction de la différence de qualité visuelle entre CRT et LCD basée sur le flou de mouvement	161
9.1	Introduction	161
9.2	Estimation locale du flou de mouvement perçu	162
9.2.1	Estimation de mouvement	163
9.2.2	Classification spatiale	164
9.2.3	Calcul du flou de mouvement	166
9.3	Prédiction de la différence de qualité visuelle entre CRT et LCD	167
9.3.1	Cumul des valeurs locales de flou de mouvement perçu	167
9.3.2	Prédiction de la différence de qualité ΔQ	170
9.3.3	Résultats	171
9.4	Observations et discussion	172
9.5	Conclusion	174
	Conclusion et perspectives	177
	Annexes	185
A	Propriétés utiles du système visuel humain	185
A.1	Introduction	185
A.2	Physiologie du système visuel humain	186
A.2.1	L'œil	186
A.2.2	La rétine	187
A.3	Adaptation du système visuel humain à son environnement	188
A.3.1	Sensibilité et adaptation à la lumière	189
A.3.2	Perception des couleurs	191
A.3.3	Mouvements de l'œil	194
A.4	Limitations du système visuel humain	195

A.4.1	Résolution spatiale et acuité visuelle	195
A.4.2	Résolution temporelle	201
A.5	Sensibilité au contraste et aux textures	203
A.5.1	Loi Weber-Fechner	203
A.5.2	Fonctions de sensibilité aux contrastes (CSF)	204
A.6	Perception du mouvement	207
A.6.1	Généralités	207
A.6.2	Cinéma et télévision : l'illusion du mouvement	208
A.7	Conclusion	210
Bibliographie		213
Liste des publications		223

Introduction générale

Depuis les débuts de la télévision dans les années 1940 et jusqu'au début du 21^{ème} siècle, les écrans à tube cathodique (CRT) ont été la technologie prédominante pour l'affichage de l'information électronique. Tout au long de la seconde moitié du 20^{ème} siècle, la technologie des tubes cathodiques s'est améliorée et est devenue un moyen peu coûteux pour afficher les signaux de télévision ou d'ordinateur. Dans les années 1990, plus de 250 millions d'écrans CRT sont produits chaque année [OCDE 00]. Les écrans à tube cathodique possèdent cependant un défaut majeur, en l'occurrence leur taille. Leur profondeur est en effet environ égale à la diagonale de l'écran (un peu moins grâce aux dernières améliorations comme le « tube plat »). De plus, ces appareils consomment énormément d'énergie. Ces caractéristiques empêchèrent donc leur utilisation dans des appareils mobiles. Dans les années 1990, la démocratisation des ordinateurs portables, assistants personnels (PDA), et autres téléphones portables, a donc nécessité le développement en masse d'un autre type d'affichage : les écrans plats. Ce sont les afficheurs à cristaux liquides (LCD), déjà utilisés depuis la fin des années 1970 sur les montres et les calculatrices par exemple, qui furent plébiscités. Gagnant en maturité d'année en année, cette technologie commença à s'installer sur le marché des moniteurs informatiques.

À la fin du 20^{ème} siècle, les écrans à tube cathodique sont encore les seuls à occuper le marché des téléviseurs, malgré la présence de quelques modèles plasma ou LCD à des prix très élevés. Cependant, les années 2000 voient l'avènement de la télévision haute-définition (TVHD), aux États-Unis, au Japon et en Australie d'abord, puis en Europe. Ce nouveau format d'image va sonner le glas des écrans à tube cathodique. En effet, la TVHD exige un écran de grande taille. Plus exactement, pour bénéficier des améliorations visuelles apportées par le format HD, la distance de visualisation recommandée est de 3 fois la hauteur de l'écran [ITU 98]. Avec ce nouveau format de diffusion, il devient donc réellement intéressant de posséder un téléviseur de grande taille. Cependant, pour des raisons de difficultés de fabrication, les écrans CRT sont limités à une taille d'écran de 100 cm en diagonale, taille pour laquelle le téléviseur est déjà un meuble imposant. Les écrans à tube cathodique ne peuvent donc être utilisés pour des applications à grands écrans. Les téléviseurs à écrans plats (LCD et plasma) envahissent alors le marché et réussissent en quelques années à faire oublier l'écran CRT pourtant centenaire. L'écran à cristaux liquides, principalement pour des raisons de coût, prendra rapidement le pas sur son rival plasma. Fin 2007, et pour la première fois dans l'histoire de la télévision, les téléviseurs à tube cathodique ne sont plus les plus vendus au monde. Dans les années qui suivent, la plupart des constructeurs annoncent la fermeture de leurs chaînes de production d'écrans CRT.

Aujourd'hui, les écrans à cristaux liquides représentent plus de 85% des ventes de téléviseurs et de moniteurs informatiques dans les pays développés [UKDisplay 09]. Cette technologie reste cependant une technologie nouvelle vis à vis de la technologie d'affichage à tube cathodique façonnée par un siècle d'évolutions adaptées à la télévision. Grâce aux nombreuses améliorations dont ils ont fait l'objet ces dernières années, les écrans LCD ont atteint une qualité très satisfaisante en terme de reproduction de couleur, de luminosité et de contraste, dépassant même les meilleurs écrans CRT dans ces domaines. Cependant, de nombreux efforts restent à fournir concernant l'affichage de contenus animés, et notamment des signaux de télévision. De part leur nature, les afficheurs à cristaux liquides sont adaptés à l'affichage de contenus fixes et la reproduction du mouvement notamment reste problématique. À ce titre, la technologie LCD ne peut être considérée comme une technologie pleinement adaptée à la télévision et ses caractéristiques doivent encore être étudiées et modifiées en vue d'égaler un jour la qualité d'affichage d'un écran CRT.

Caractérisation objective : évaluer la technologie via des mesures instrumentales du flou de mouvement

Dès le début de leur commercialisation en tant que moniteurs informatiques, les écrans à cristaux liquides ont été pointés du doigt pour leurs caractéristiques temporelles jugées faibles. On parlait alors de « rémanence » ou de « ghost effect » et nombreux étaient ceux qui préféraient utiliser un écran à tube cathodique, notamment pour les applications vidéo-ludiques. Le temps de réponse des pixels de l'écran, d'une valeur moyenne de 25 millisecondes à cette époque, était considéré à juste titre comme le responsable de la mauvaise reproduction du mouvement sur les écrans LCD. Les différents constructeurs essayèrent donc par tous les moyens possibles de réduire ce temps de réponse. Tout naturellement, cette mesure devint donc la caractéristique principale de l'écran lors de sa commercialisation. Tous les mois, un nouveau modèle voyait le jour avec un temps de réponse inférieur que les précédents, jusqu'à atteindre des temps de réponse de l'ordre de 2 millisecondes à partir de 2007. Pourtant, même avec des temps de réponse très faibles la qualité d'affichage de certains écrans reste très moyenne. Et pour cause, même avec un temps de réponse nul, un « flou de mouvement » reste présent. Ce défaut particulier des écrans à cristaux liquides a été identifié dès le début des années 2000 [Kurita 01]. Ses deux principales causes sont *l'affichage par maintien* caractéristique de ce type d'écran et *le comportement du système visuel humain* face à des objets en mouvement.

Plusieurs problèmes se posent quant à l'utilisation de la mesure du temps de réponse pour caractériser la qualité de la reproduction du mouvement sur les écrans à cristaux liquides. Tout d'abord, il faut savoir qu'il n'existe pas de consensus concernant le protocole de mesure du temps de réponse. Il n'y a pas une valeur de temps de réponse par écran mais 255×256 valeurs (dans le cas d'un écran 8 bits), c'est à dire une valeur différente pour chaque transition gris-à-gris de l'écran. Si différentes recommandations ont été émises concernant la valeur à faire apparaître dans les caractéristiques techniques de l'écran (temps de réponse du noir vers le blanc et vice-versa, moyenne des temps de réponse gris-à-gris d'un ensemble de transitions choisies, etc.), force est de constater que chaque constructeur utilise une méthode qui lui est propre, en s'arrangeant pour qu'elle donne une valeur finale la plus faible possible. De plus, ce n'est pas le temps de réponse des pixels le principal res-

ponsable du flou de mouvement, mais bien la forme de la réponse. On parle alors d’affichage par maintien sur les écrans LCD, par opposition à l’affichage par impulsion des écrans à tube cathodique. En fait, on estime à 20% la contribution du temps de réponse à l’apparition du flou de mouvement sur les écrans LCD [Pan 05]. Pourtant, encore aujourd’hui, les constructeurs continuent d’afficher le temps de réponse comme la caractéristique qui permettrait d’évaluer la qualité de la reproduction du mouvement sur leurs produits. Cette mauvaise habitude est désignée en anglais par le terme « specsmanship » et le temps de réponse en constitue un des plus beaux exemples qui soient. Il est en réalité impossible de prédire la quantité de flou de mouvement présente sur un écran LCD sur la seule base de son temps de réponse. Un écran avec un temps de réponse affiché de 20 ms pourrait bien s’avérer au final de meilleure qualité qu’un écran affichant 2 ms, pour la plus grande confusion des utilisateurs.

Depuis plusieurs années, le flou de mouvement des écrans LCD constitue le cheval de bataille des constructeurs d’écrans. Ce problème qui pouvait sembler marginal sur les moniteurs informatiques, est devenu crucial sur les téléviseurs destinés à afficher des contenus en mouvement. De nombreuses techniques, matérielles ou logicielles, ont été développées dans le but de réduire ce défaut. Cependant, en l’absence d’une mesure précise et claire du flou de mouvement, il est difficile d’évaluer exactement les améliorations apportées par ces dispositifs. Il est également impossible de comparer deux modèles de téléviseurs, chacun proposant des solutions différentes à ce problème. Il existe donc un besoin clairement défini : la mise en place d’un protocole précis permettant de mesurer le flou de mouvement sur LCD et la définition d’une mesure globale et objective autorisant la comparaison des écrans entre eux en terme de qualité de reproduction du mouvement. Plusieurs dispositifs de mesure du flou de mouvement existent à l’heure actuelle. La plupart reproduisent le comportement de l’œil humain à l’aide d’une caméra rapide : la caméra est synchronisée à l’écran et déplacée de manière à suivre un contour en mouvement. Néanmoins, ce protocole de mesure est coûteux et complexe à mettre en œuvre. Les différences entre capteurs, les conditions de mesure, peuvent varier d’un laboratoire à l’autre et interdisent la comparaison des mesures effectuées. En conséquence, il n’existe pas à l’heure actuelle de consensus entre les différents constructeurs concernant la publication d’un protocole standard de mesure du flou de mouvement. Le développement d’un protocole de mesure simple à mettre en œuvre et facilement reproductible d’un laboratoire à l’autre pourrait permettre cette standardisation.

Modélisation psychovisuelle : évaluer l’impact du flou de mouvement sur la qualité des images affichées

L’évaluation du flou de mouvement doit également prendre en compte les caractéristiques de la vision humaine. En effet, le flou de mouvement trouve son origine dans le fonctionnement intrinsèque de la technologie d’affichage à cristaux liquides, mais également et surtout dans le comportement du système visuel humain. En réalité, le flou de mouvement n’est pas présent sur l’écran, mais seulement sur le capteur qui suit le mouvement des objets en déplacement sur cet écran, c’est à dire sur la rétine dans le cas d’un observateur humain. Il ne suffit donc pas de caractériser cette distorsion de l’image de manière technologique, à l’aide d’instruments optiques. Il convient également d’étudier la réponse du système visuel au flou de mouvement, pour connaître la relation entre le flou mesuré et

le flou réellement perçu par les observateurs. Par exemple, le flou de mouvement mesuré de manière objective est généralement le plus important pour les contours de faible amplitude lumineuse. En effet, dans ces conditions ce sont les transitions gris-à-gris les moins réactives qui sont sollicitées. Mais ce sont également sur ces contours de faible amplitude que le flou de mouvement sera le plus difficilement perçu par les observateurs. On peut donc supposer l'existence d'une différence importante entre le flou de mouvement objectif et le flou de mouvement perçu, variable selon les transitions considérées. La caractérisation de la perception du flou de mouvement doit permettre de quantifier cette différence.

La perception visuelle est un sujet vaste et passionnant, largement étudié au cours des dernières décennies. De nombreux modèles du système visuel humain ont été développés sur des aspects tels que la perception du contraste, des couleurs, ou encore du mouvement. L'utilisation de modèles existants pour caractériser la perception du flou de mouvement est cependant délicate étant donné la complexité de ce stimulus spatio-temporel. De nombreux paramètres entrent en jeu, tels que le mouvement rétinien, la vitesse de ce mouvement, la luminance et le contraste de l'objet en mouvement, etc. Dans ce contexte particulier, il semble donc plus prudent d'étudier directement le flou de mouvement que de vouloir prédire la réponse du système visuel à partir de modèles existants. Il existe néanmoins peu de travaux sur ce sujet. La difficulté à définir un protocole de test dans lequel l'observateur suit des yeux un objet en mouvement tout en évaluant la quantité de flou perçu est certainement une des raisons principales de ce manque. De même, pour être rigoureux, ces travaux doivent être effectués sur un ensemble assez large de transitions gris-à-gris et sur plusieurs écrans, ce qui implique une quantité de travail conséquente.

La modélisation de la perception du flou de mouvement est de première importance pour établir une mesure *perceptive* du flou de mouvement. C'est à dire une mesure basée sur la quantité de flou de mouvement effectivement perceptible sur l'écran et non pas seulement sur le flou de mouvement mesuré à l'aide d'instruments optiques. Il reste cependant une interrogation concernant l'impact du flou de mouvement sur la qualité visuelle d'un contenu affiché à l'écran. Dès 2005, une étude de l'Union International des Télécommunications [ITU 05] pointait les différences de qualité entre CRT et LCD, en indiquant que le flou de mouvement était le défaut jugé le plus gênant par les observateurs pour les séquences contenant de forts mouvement. L'étude de la perception du flou de mouvement sur des stimuli simples tels que des contours en mouvement doit donc être complétée avec des stimuli naturels, et la contribution du flou de mouvement à la qualité finale d'un contenu identifiée. Les améliorations, matérielles ou logicielles, apportées par les constructeurs pour réduire le flou de mouvement pourront alors être évaluées sur des contenus de types télévisuels, à travers le prisme de la vision.

Objectifs de la thèse

Les objectifs de la thèse se déclinent suivant trois grands axes.

- Le premier axe concerne la mise en œuvre d'un protocole de mesure du flou de mouvement. L'objectif est de définir une méthode simple et efficace, dont la mise en œuvre soit reproductible dans des conditions similaires d'un laboratoire à l'autre. Les résultats obtenus avec cette nouvelle méthode devront être comparables à ceux des méthodes existantes.

- Le deuxième axe concerne la modélisation psychovisuelle de la perception du flou de mouvement. L'objectif est la définition d'une mesure perceptive du flou de mouvement. C'est à dire une mesure qui tiennent compte de l'aspect non linéaire de la perception visuelle dans le calcul d'une valeur globale de flou de mouvement. Cette étude psychovisuelle s'intéressera également à la gêne visuelle provoquée par le flou de mouvement pour différents types de contenus.
- Enfin, le troisième axe concerne l'impact du flou de mouvement sur la qualité globale des séquences vidéo affichées sur un écran à cristaux liquides. Par l'intermédiaire d'une estimation locale du flou de mouvement perçue basée sur les modèles précédents, l'objectif est de prédire la perte de qualité visuelle sur LCD mise en évidence par une campagne de tests d'évaluation subjective de la qualité visuelle.

Organisation du mémoire

Ce mémoire de thèse est organisé en quatre parties.

La première partie décrit le contexte de ces travaux et pose les problématiques liées aux écrans à cristaux liquides. Le chapitre 1 de ce mémoire traite des notions générales concernant l'affichage et la télévision, s'intéressant aux évolutions technologiques qui ont permis au fil des années de constamment améliorer la qualité d'usage. L'écran à tube cathodique est décrit et ses principales limitations identifiées, expliquant l'émergence de nouvelles technologies d'affichage depuis une dizaine d'années. Le chapitre 2 s'intéresse plus précisément au fonctionnement des écrans à cristaux liquides, et à la formation des images sur ce type d'écran. Cette nouvelle technologie d'affichage est très différente de l'écran à tube cathodique utilisé jusque là en télévision et souffre de nouvelles distorsions visuelles, inexistantes sur CRT. Ces nouveaux défauts créent le besoin de nouvelles méthodes d'identification et de caractérisation.

La seconde partie traite de la mesure objective du flou de mouvement par l'intermédiaire d'instruments électroniques. Le chapitre 4 présente une description du flou de mouvement et de son origine à la fois technologique (affichage par maintien) et perceptive (mouvement de poursuite). Un état de l'art est fait des techniques de mesures temporelles sur écran LCD. Dans le chapitre 5, nous présentons une méthode analytique de mesure du flou de mouvement à partir de la réponse indicielle des pixels de l'écran. Cette méthode est décrite et les résultats obtenus sont comparés avec ceux obtenus par la méthode « classique » utilisant une caméra rapide. Les avantages et défauts des deux méthodes sont confrontés.

La troisième partie traite de la caractérisation psychovisuelle du flou de mouvement et de sa modélisation. Le chapitre 6 décrit l'étude de l'influence du mouvement rétinien sur la perception du contraste, à des niveaux supra-limaires (au delà du seuil différentiel de visibilité). Dans le chapitre 7, la perception du flou de mouvement est étudiée via un protocole original d'égalisation de flou. Ces expériences permettent de mesurer de la quantité de flou réellement perçue par les observateurs sur un contour en mouvement, pour différentes vitesses, différentes dynamiques lumineuses, et sur plusieurs écrans. Enfin, le chapitre 8 présente les expérimentations psychophysiques menées sur des stimuli plus complexes (textes défilants, images en mouvement, séquences synthétiques) dans le but d'étudier la gêne associée à la présence de flou de mouvement.

Enfin, dans la dernière partie, les résultats obtenus sont mis en relation avec l'évaluation subjective de la qualité sur écrans à cristaux liquides. Le chapitre 9 présente les tests d'évaluation subjective de la qualité visuelle, menés sur écran CRT et sur écran LCD. Dans le chapitre 10, la perte de qualité observée sur les écrans LCD dans le cas de contenus en mouvement est mise en relation avec la quantité de mouvement dans la séquence. Un modèle de prédiction de cette différence de qualité est conçu, basé sur l'estimation locale du flou de mouvement à partir des résultats de la partie 3.

Première partie

Visualisation d'images animées, du CRT au LCD, concepts technologiques et nouveaux enjeux

Chapitre 1

Systèmes d’affichage électronique, évolution et limites du CRT

Sommaire

1.1	Introduction	9
1.2	Concepts de base	10
1.2.1	Structure d’un système d’affichage	10
1.2.2	Notions et terminologie	11
1.2.3	Distinction entre écran et contenu	13
1.3	Télévision et écran à tube cathodique	13
1.3.1	Naissance de la télévision	13
1.3.2	Écran à tube cathodique (CRT)	14
1.3.3	Standardisation de la télévision analogique	16
1.4	Vers une amélioration de la qualité d’usage	16
1.4.1	Immersion, impact et expérience comparable au cinéma	16
1.4.2	Limites de l’écran à tube cathodique et de la télévision standard	17
1.4.3	Conditions pour une qualité d’usage accrue	18
1.4.4	Nouveaux écrans	19
1.5	Conclusion	21

1.1 Introduction

Un système d’affichage électronique est un dispositif qui présente une information visuelle, générée à partir d’un signal d’entrée électrique, soit par émission de lumière, soit par modulation d’une source lumineuse existante. Les premiers systèmes d’affichage électronique apparaissent lors de l’avènement de la télévision au milieu du 20^{ème} siècle. Les téléviseurs de l’époque sont constitués d’un tube cathodique par l’intermédiaire duquel un faisceau d’électrons vient exciter les luminophores présents sur l’écran. Soixante ans plus tard, le principe de fonctionnement des écrans à tube cathodique n’a pas changé. Cependant, après un demi-siècle d’améliorations et de perfectionnements, l’écran à tube

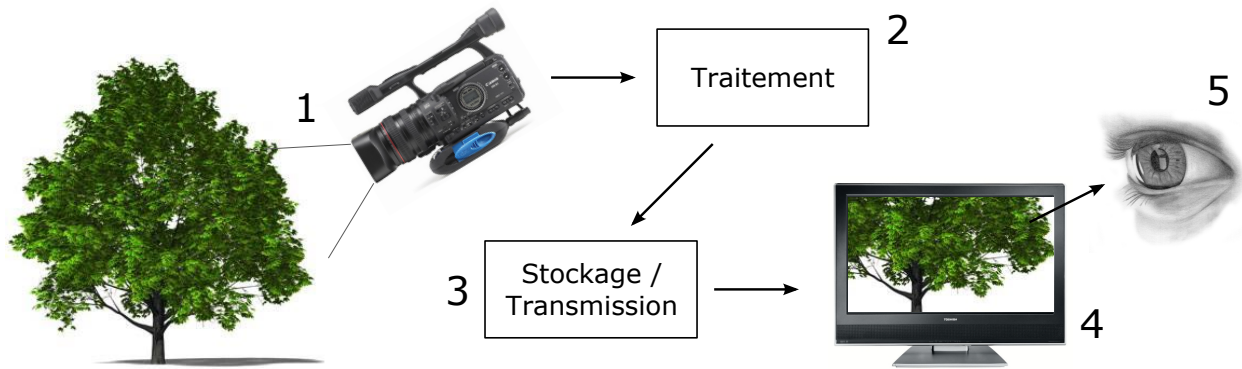


FIGURE 1.1 – Les éléments d'un système d'affichage sont : (1) l'image source, ici un objet réel vu par un système d'acquisition (une caméra), (2) le traitement de l'image, (3) le stockage et/ou la transmission de l'image, (4) l'affichage proprement dit, et (5) l'observateur.

cathodique a atteint ses limites, ouvrant la voie à de nouveaux systèmes d'affichage électronique au fonctionnement radicalement différent.

Dans ce premier chapitre, les concepts relatifs aux systèmes d'affichage électronique sont décrits de manière générale. Les différentes évolutions de la télévision sont ensuite rappelées, de manière à comprendre les raisons qui ont provoqué la fin des écrans à tube cathodique. Finalement, nous ferons le point sur les nouvelles technologies et sur les enjeux à venir concernant l'affichage électronique.

1.2 Concepts de base

1.2.1 Structure d'un système d'affichage

La chaîne complète d'un système d'affichage peut être divisée en quelques éléments de base (figure 1.1), présents dans n'importe quelle application sous une forme ou une autre.

L'image source peut prendre de nombreuses formes : une scène naturelle vue à travers un capteur optique (appareil photographique, caméra), ou encore une information structurée traduite sous une forme visuelle (texte, diagrammes, etc.).

Le traitement de l'image est nécessaire pour convertir l'information brute issue de la source dans une forme qui soit adaptée au reste du système (et plus particulièrement au système d'affichage). La « correction gamma » est un exemple de traitement simple effectué en amont d'un système de diffusion pour la télévision. En synthèse d'images, c'est le « rendu » qui consiste à écrire l'image dans la mémoire tampon sous une forme affichable (conversion de la représentation 3D d'une scène vers une image 2D par exemple).

La transmission de l'image a pour but d'acheminer l'information jusqu'à l'écran. Les images peuvent être transmises au fur et à mesure de leur acquisition. Elles peuvent aussi être stockées en prévision de leur affichage. Le stockage peut prendre diverses formes selon les applications : enregistrement

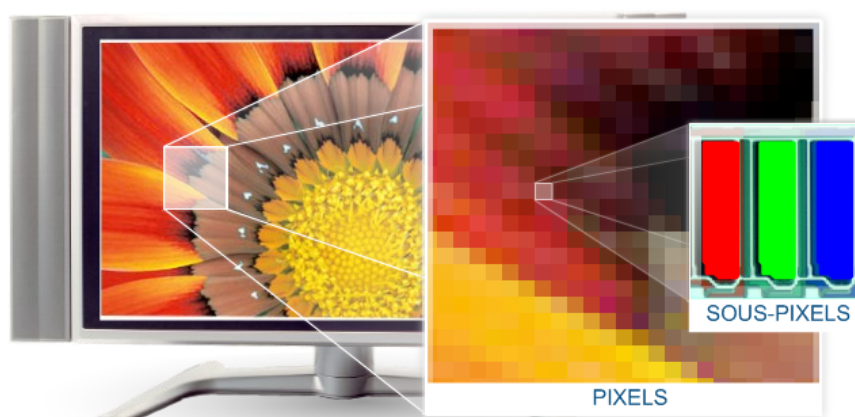


FIGURE 1.2 – Affichage d’une image sur un écran.

sur un DVD, dans la mémoire d’un appareil photo, ou encore dans la mémoire tampon d’un système de synthèse d’images. La zone de stockage d’un système d’affichage peut servir à différer la présentation des images à l’observateur. C’est aussi souvent à ce niveau qu’une image formée de plusieurs composantes est reconstruite et formatée de manière à constituer l’image finale vue par l’observateur.

Dans ce mémoire de thèse, nous nous focalisons sur les deux derniers éléments, à savoir l’écran et l’observateur. Ce dernier est certainement la partie la plus importante de la chaîne puisque c’est lui qui va déterminer les pré-requis du système d’affichage en termes de performances et de qualité. Il est important de garder à l’esprit qu’idéalement un système d’affichage doit répondre aux besoins de la vision humaine en s’appuyant sur les limitations de celle-ci.

1.2.2 Notions et terminologie

Dans cette section, les notions de base concernant les écrans et le signal qu’ils affichent sont décrites indépendamment de la technologie utilisée. Il s’agit d’expliquer de manière concise comment une image est affichée, et de définir les différents termes qui seront utilisés dans la suite du mémoire.

1.2.2.1 Pixels et image

Un écran peut être vu comme une matrice de points lumineux, désignés par le terme « pixel » (de l’anglais *picture element*). C’est la juxtaposition de plusieurs milliers (voire millions) de ces pixels qui forme l’image. Chaque pixel constitue donc un point de l’image, défini par sa position et par sa valeur (couleur). La couleur d’un pixel est créée à partir de trois couleurs primaires, généralement le rouge, le vert et le bleu. Un pixel est donc en réalité constitué de trois sous-pixels (cf. figure 1.2) qui émettent chacun une des couleurs primaires avec plus ou moins d’intensité. La combinaison des trois couleurs primaires à différents degrés d’intensité permet de reproduire une couleur donnée, grâce au caractère trichromate du système visuel humain (cf. section A.3.2).

Il convient ici de différencier le pixel d'une image et le pixel d'un écran. Le pixel d'une image est un élément ponctuel sans dimension qui correspond à un échantillon discret de l'image originale (continue). Le pixel d'un écran a une dimension physique donnée. Cette dimension est d'importance et ne doit pas être perçue par l'observateur sous peine de briser l'illusion de continuité spatiale de l'image affichée.

1.2.2.2 Résolution, taille et format

La résolution d'un écran est définie comme le nombre de pixels qu'il possède. Il est généralement exprimé par le nombre de pixels contenus dans une ligne suivi du nombre de pixels contenu dans une colonne (ou bien le nombre de colonnes suivi par le nombre de lignes, ce qui revient au même). Par exemple, un écran avec une résolution de 1920×1080 pixels possède 1080 lignes de 1920 pixels chacune. Une autre caractéristique importante est la taille de l'écran, c'est à dire sa dimension physique. Celle-ci est généralement mesurée sur une diagonale de l'écran et exprimée en pouces¹. Les dimensions de l'écran, pour une diagonale donnée, dépendent alors du rapport de la largeur sur la hauteur. Celui-ci est généralement représenté par deux nombres entiers plutôt que par une valeur décimale. Par exemple, 4:3 et 16:9 sont les rapports les plus fréquents pour les télévisions ; on rencontre également les rapports 5:4 et 16:10 pour les moniteurs informatiques.

Dans le cas d'une image, on désigne par « format » le nombre de pixels. À titre d'exemple, le format CIF (de l'anglais *Common Intermediate Format*) correspond à 352×288 pixels. Comme pour la résolution d'un écran, le nombre de pixels dans la dimension horizontale (nombre de colonnes) est donné avant le nombre de pixels dans la dimension verticale (nombre de lignes). Dans la littérature, on rencontre également le terme « résolution » pour désigner le format d'une image.

L'emploi du terme « résolution », dans un cas comme dans l'autre, est inapproprié. La résolution d'une image s'exprime en pixels par pouce (en anglais, *dots per inch*, dpi) et correspond au rapport entre le nombre de pixels et la taille physique de l'image. Dans le cas d'un écran, la résolution au sens premier est désignée par le terme *pitch*, qui correspond à la taille d'un pixel et s'exprime en millimètre par pixel.

Dans la suite de ce mémoire nous parlerons cependant de « résolution » pour désigner le nombre de pixels d'un écran, car le terme n'entre pas en conflit avec le pitch et qu'il est largement utilisé. La « taille » d'un écran désignera toujours sa diagonale, sauf mention contraire. Enfin, on préférera parler de « format » pour spécifier le nombre de pixels d'une image ou, plus généralement, le format d'une vidéo (cf. paragraphe suivant).

1.2.2.3 Fréquence d'affichage et débit d'images

La fréquence d'affichage, ou fréquence de rafraîchissement, d'un écran correspond au nombre de fois que l'écran est mis à jour dans un laps de temps d'une seconde. Elle s'exprime en hertz. On utilise parfois l'écriture $1920 \times 1200@60\text{Hz}$ pour spécifier à la fois la résolution et la fréquence de rafraîchissement d'un écran.

1. Le pouce technique international vaut 2.54 centimètres, son abréviation est in ou ”.

Le débit d'images, ou cadence d'images, d'une séquence vidéo correspond à la vitesse à laquelle doivent être présentées les images de la séquence. Elle s'exprime en nombre d'images par seconde, que l'on trouve parfois abrégé en i/s, ou encore fps (de l'anglais *frame per second*). Les formats de diffusion de la télévision sont généralement spécifiés en mentionnant le nombre de lignes d'une image suivi de la cadence d'images. Par exemple, les formats de diffusion haute-définition sont le 1080/50i, le 720/50p et 1080/50p. Les lettres *i* et *p* précisent si le format est entrelacé (en anglais *interlaced*) ou progressif (cf. section 1.3.2.2).

1.2.3 Distinction entre écran et contenu

Il est important de lever toute ambiguïté entre les termes relatifs à l'écran et les termes relatifs au contenu affiché sur celui-ci.

Les écrans numériques (ou plus précisément les « écrans à pixels fixes ») tels que les écrans à cristaux liquides ou les écrans plasma ont une résolution dite « native ». Quelle que soit l'image qu'ils reçoivent en entrée, elle sera affichée à la résolution native. Si le format d'entrée est différent de la résolution native de l'écran, l'image devra être transformée de manière à s'adapter à la mosaïque de pixels de l'écran. Si l'image d'entrée est plus grande, elle sera réduite via un processus de filtrage et de sous-échantillonnage. Si elle est plus petite, elle sera sur-échantillonnée (bien qu'elle puisse également être affichée sans transformation sur une portion limitée de l'écran). Les opérations de sous- et sur-échantillonnage sont effectuées en temps réel par l'écran. Elles ne sont pas simples, particulièrement lorsque le facteur d'échelle n'est pas entier, et introduisent des défauts visuels plus ou moins importants selon la méthode utilisée. Par exemple, utiliser un écran « HD ready » d'une résolution native de 1366×768 pixels pour afficher la télévision au format standard (576i) n'apportera pas d'amélioration de qualité, et risque même d'introduire des dégradations.

Pour les écrans à tube cathodique (analogique), le problème est différent. S'ils possèdent bien une résolution native, aucune opération de sur- ou de sous-échantillonnage n'est nécessaire pour afficher une image d'un format différent. La mise à l'échelle de l'image affichée est inhérente à la technologie utilisée dans les écrans à tube cathodique.

1.3 Télévision et écran à tube cathodique

1.3.1 Naissance de la télévision

La naissance de la télévision fut le fruit d'un lent travail d'amélioration collective entre chercheurs de différents pays, à la fin du 19^{ème} siècle et au début du 20^{ème} siècle. Les premiers systèmes de télévision étaient basés sur le disque de Paul Nipkow [Nipkow 84] : un disque tournant percé de trous, permettant de « scanner » la scène point par point et ligne par ligne. L'affichage utilisait un système similaire mais inversé. La conversion d'une onde lumineuse en impulsion électronique permettant la transmission de l'image fut rendu possible par la découverte de l'effet photo-électrique par Smith en 1873 [Smith 73]. Ces systèmes électro-mécaniques furent utilisés au début du 20^{ème} siècle, un système complet de diffusion, le *Scophony*, fut même développé dans les années 1930 en Grande-Bretagne [Hennessy 05].

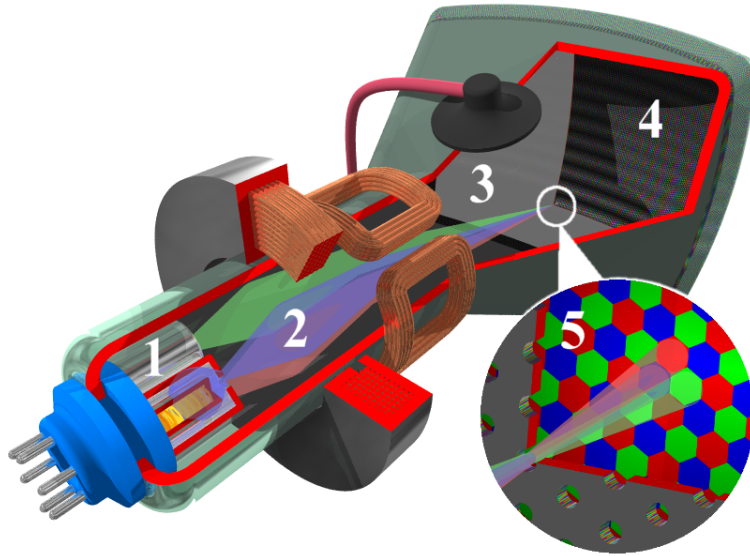


FIGURE 1.3 – Écran à tube cathodique. 1 : canons à électrons. 2 : faisceaux d'électrons. 3 : masque permettant de guider chacun des trois faisceaux vers le luminophore de la couleur correspondante. 4 : face interne de l'écran, où sont situés les luminophores (pastilles phosphorescentes). 5 : gros plan sur le masque et les luminophores. (Source : Wikipédia.)

La télévision électro-mécanique fut cependant très vite remplacée par la télévision électronique, celle que nous connaissons aujourd'hui. L'essor de la télévision tout-électronique est dû au développement du tube cathodique, utilisé aussi bien pour enregistrer l'image que pour l'afficher. La paternité de la télévision moderne est encore largement disputée aujourd'hui. Il semblerait néanmoins qu'elle revienne au hongrois Kálmán Tihanyi qui breveta dès 1926 un système de télévision électronique. Dans les années qui suivirent, des systèmes plus complets furent développés. On peut citer par exemple l'*Iconoscope* de Vladimir Zworykin [Zworykin 35] et l'*Image Dissector* de Philo Farnsworth [Farnsworth 37]. Dès les années 1940, un système combinant les qualités de l'iconoscope et de l'image dissector voit le jour : l'image-orthicon. Il sera utilisé jusqu'à la fin des années 1960 avant d'être remplacé par des systèmes plus modernes mais fonctionnant sur le même principe tels que le Vidicon, le Plumbicon ou le Trinicon [Abramson 03].

1.3.2 Écran à tube cathodique (CRT)

Le système de télévision tel qu'on le connaît est donc intimement lié au tube cathodique (CRT, de l'anglais *Cathod Ray Tube*). En fait, l'un n'aurait pas vu le jour sans l'autre et vice-versa. C'est l'utilisation des tubes cathodiques qui a rendu la télévision possible et c'est la télévision qui a justifié l'intérêt pour les tubes cathodiques.

1.3.2.1 Fonctionnement

La figure 1.3 présente la structure d'un écran CRT. Le tube cathodique est constitué d'une cathode et d'une anode (l'écran). Lorsqu'on applique un fort champ électrique entre l'anode et la cathode,



FIGURE 1.4 – Illustration de l'entrelacement. (a) Image originale. (b) Image affichée par un balayage progressif de l'écran. (c) Image affichée par un balayage entrelacé. La différence de luminosité est moins gênante dans le cas d'un affichage entrelacé, la finesse des lignes permet à l'observateur de ne pas percevoir le clignotement des lignes. Dans le cas d'un affichage progressif, le clignotement est perceptible car il implique une grande partie de l'image.

les électrons sont extraits de la cathode, ce qui provoque l'apparition d'un flux d'électrons en direction de l'anode. Ce flux à très grande vitesse est focalisé électro-magnétiquement par une grille. L'ensemble forme le canon à électron. Le flux focalisé arrive ensuite alors sur l'anode recouverte d'une matière phosphorescente : les luminophores. Quand les électrons frappent cette substance un flux lumineux est émis, d'intensité proportionnelle à celle du flux d'électrons.

L'affichage se fait par balayage : toute la face du tube est parcourue par le flux d'électron selon un trajet bien défini, et l'image est créée en faisant varier l'intensité du faisceau, et donc l'intensité du flux lumineux émis par chaque luminophore. La couleur est créée grâce à la combinaison de trois faisceaux distincts qui sont dirigés chacun vers un type de sous-pixels à l'aide d'un masque ou d'une grille.

1.3.2.2 Entrelacement

Pour des raisons de commodités techniques, il a été décidé de synchroniser la fréquence d'affichage des écrans CRT sur la fréquence du réseau électrique (60 Hz aux États-Unis, 50 Hz en Europe). Cependant les moyens techniques de l'époque ne permettait pas d'atteindre une telle cadence d'images pour les formats choisis (autour de 500 lignes), aussi bien au niveau de l'enregistrement que de l'affichage. La cadence d'images des formats utilisés a donc été fixé à 30 i/s aux États-Unis et à 25 i/s en Europe. Ce qui était d'autant plus pratique que la cadence d'images du cinéma était de 24 i/s, donc il était aisé d'adapter le format pour un affichage à la télévision.

Cependant, au cinéma l'image entière est affichée en une seule fois, par projection. Dans ces conditions, l'observateur ne perçoit pas de clignotement. Sur un écran CRT, l'affichage se fait par balayage, point par point. Par conséquent, au moment où le bas de l'image est affiché (en pleine luminosité), la luminosité du haut de l'image a déjà diminuée, d'autant plus qu'avec les luminophores de l'époque,

la luminosité décroissait rapidement (cf. figure 1.4b) . Dans ces conditions, le clignotement de l'écran était perceptible. Il a donc fallu trouver un moyen de contourner le problème : l'entrelacement.

Au lieu d'être affichée progressivement, ligne après ligne, l'image est affichée en deux temps : d'abord les lignes paires, puis les lignes impaires (cf figure 1.4c), avec une cadence de 60 ou 50 demi-images (ou trames) par seconde. De cette façon, la différence de luminosité due aux phosphores est répartie sur l'ensemble de l'image. Étant donné la taille très fine d'une ligne, le clignotement n'est plus perceptible. On obtient donc artificiellement une cadence de 60 ou 50 i/s mais le nombre de lignes de chaque trame est divisé par deux.

Pour résumer, l'entrelacement est donc une technique qui permet d'économiser le débit d'images, en évitant le clignotement, le tout sans détériorer la reproductibilité du mouvement.

1.3.3 Standardisation de la télévision analogique

A partir des années 1950 et 1960, les formats de diffusion de la télévision vont être progressivement standardisés dans tous les pays du globe. Cette standardisation coïncide plus ou moins avec l'introduction de la télévision couleur. On distingue trois standards de télévision couleur sur la planète (cf. figure 1.5) :

- Le système NTSC (pour *National Television System Committee*), aux États-Unis dès 1953, est composé de 525 lignes (480 ligne utiles) et cadencé à 30 images entrelacées par seconde (donc 60 trames par seconde).
- Le système SECAM (pour *Séquentiel Couleur À Mémoire*), en France en 1967, est composé de 625 lignes (576 lignes utiles) et cadencé à 25 images entrelacées par seconde (donc 50 trames par seconde).
- Le système PAL (pour *Phase Alternating Line*), au Royaume-Uni et en Allemagne de l'Ouest en 1967, est composé, comme le SECAM, de 625 lignes (576 lignes utiles) et cadencé à 25 images entrelacées par seconde (donc 50 trames par seconde).

La fréquence du réseau électrique de chaque pays est utilisée à des fins de synchronisation, ce qui explique la différence dans les cadences d'images des trois standard (60 Hz au États-Unis et au Japon, 50 Hz en Europe).

1.4 Vers une amélioration de la qualité d'usage

1.4.1 Immersion, impact et expérience comparable au cinéma

Depuis ses débuts, la télévision a cherché à s'approcher des conditions de visualisation du cinéma, de manière à offrir une qualité d'usage comparable. En effet, le « petit écran » a toujours souffert d'un manque d'immersion et de profondeur par rapport à son grand frère. Les principales différences entre la télévision et le cinéma sont la taille de l'écran et la résolution des images affichées.

La taille de l'écran dans le champ visuel est un facteur déterminant dans l'immersion et l'impact ressentis par l'observateur. La figure 1.6 illustre la notion de champ visuel. Le champ visuel humain

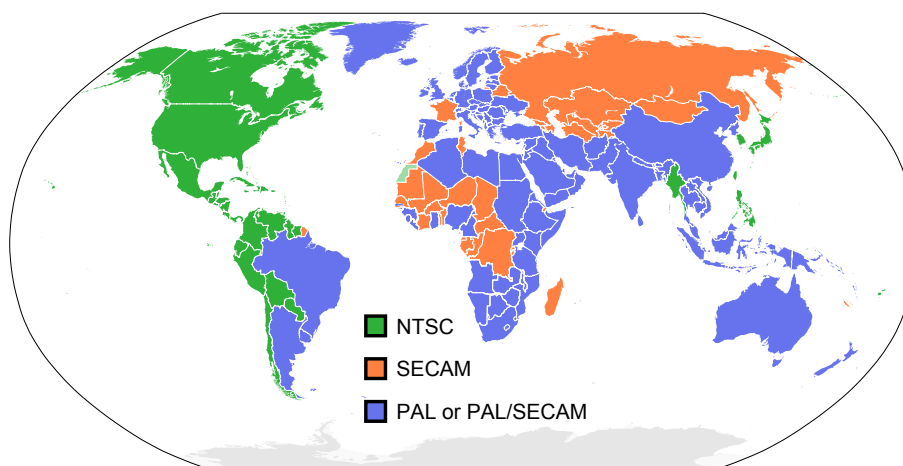


FIGURE 1.5 – Répartition des trois standard de la télévision couleur.

couvre environ 200 degrés visuels en largeur par 135 degrés visuels en hauteur [Wandell 95]. Plusieurs études japonaises datant du début des années 1980 [Fujio 82, Yuyama 82] ont cherché à déterminer les conditions d’affichage à partir desquelles l’observateur se sent immergé dans le contenu qu’il regarde. À partir de nombreuses expériences subjectives, les auteurs ont déterminé que pour obtenir un effet comparable à celui du cinéma, il faut que l’écran de télévision soit vu sous un angle d’environ 20 à 25 degrés en hauteur au minimum, soit environ 2.5 fois la hauteur de l’écran ($D = 2.5H$). Cependant, la « distance de confort », c’est à dire la distance moyenne à laquelle les téléspectateurs se placent, ne dépend pas simplement de la taille de l’image. De nombreux autres paramètres entre en jeu : disposition du salon, nombre de personnes devant l’écran au même moment, etc. Cette distance de confort a été mesurée pour différentes résolutions et pour différentes tailles d’images, dans des conditions d’utilisation normales, elle se situe généralement entre 2 m et 2.5 m. [Yuyama 82]. La principale conséquence des résultats de ces études est que, à cette distance d’observation, la hauteur d’un écran de télévision doit être de l’ordre de 1 mètre pour que les observateurs ressentent pleinement une expérience comparable au cinéma. Il faut donc chercher à s’approcher au mieux de ces conditions idéales pour améliorer encore la qualité d’usage des observateurs.

1.4.2 Limites de l’écran à tube cathodique et de la télévision standard

La taille de l’écran doit donc être augmentée, et la distance de visualisation réduite, de manière à augmenter l’immersion et la profondeur ressentie lors d’une expérience télévisuelle. Deux problèmes majeurs se posent alors pour les systèmes de télévision standard (signal diffusé et écran à tube cathodique) :

- la taille physique de l’écran ;
- la résolution de l’écran et le format des images affichées.

Le premier problème est d’ordre purement technique. Les écrans CRT sont constitués d’un canon à électrons, de plusieurs bobines de grosse taille, de défecteurs et de grilles. Par conséquent, la

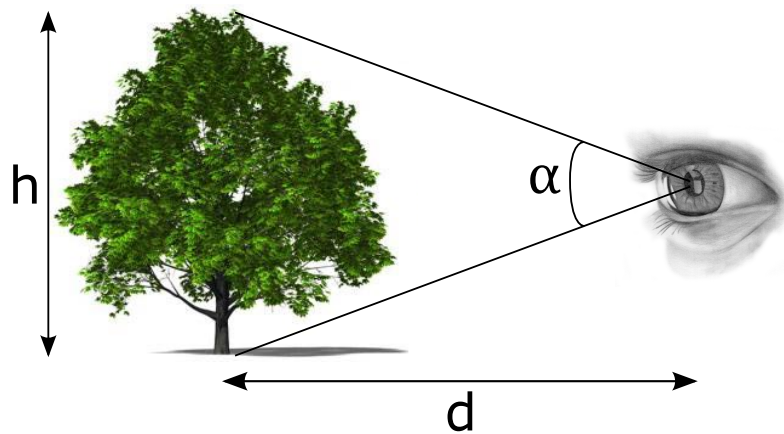


FIGURE 1.6 – Illustration de l'angle visuel. la taille d'un objet peut être exprimé en degrés visuels, c'est à dire la valeur de l'angle α sous lequel est vu l'objet. Cette valeur dépend à la fois de la taille de l'objet h et de la distance d'observation d : $\frac{h}{2d} = \tan(\frac{\alpha}{2})$.

profondeur du téléviseur est proportionnelle à la taille de l'écran proprement dit. Le principal inconvénient est alors l'encombrement de l'appareil. Au delà d'une diagonale de 30" les écrans CRT sont de véritables meubles : ils sont lourds et encombrants, et leur profondeur est proportionnelle à leur dimension. De plus, leur prix et leur consommation électrique varient dans les mêmes proportions.

Le second problème est lié au fonctionnement de l'écran CRT. Comme on l'a vu précédemment, l'écran CRT affiche les images par balayage, ligne par ligne. Le balayage crée un type de gêne spécifique la « perturbation du balayage » (en anglais *scanning line disturbance*) : en deçà d'une certaine distance d'observation, la structure des lignes devient perceptible [Yuyama 82]. Selon cette même étude, l'espace entre 2 lignes doit être inférieur à 1.2 minutes d'arc visuel pour que cette gêne n'apparaisse pas. Cette valeur est à mettre en relation avec les valeurs typiques d'acuité visuelle définies en annexe (section A.4.1.4).

Si l'on souhaite proposer un système de télévision permettant l'usage d'un écran d'une hauteur équivalente à 20 degrés visuels, pour une expérience visuelle comparable au cinéma, il faut par conséquent que le contenu affiché soit composé au minimum d'un millier de lignes.

1.4.3 Conditions pour une qualité d'usage accrue

Des deux problèmes soulevés dans la section précédente il vient deux conditions nécessaires à l'augmentation de la qualité d'usage et de l'impact :

- trouver un remplaçant au vieil affichage à tube cathodique, de manière à pouvoir augmenter la taille des téléviseurs ;
- augmenter le format des images diffusées à l'écran, pour profiter de l'immersion apportée par un grand écran sans perdre en qualité.

C'est dans cette optique que dès la fin des années 1980, de nombreux projets concernant les formats haute-définition ont vu le jour. On peut citer par exemple la norme analogique MUSE (pour *Multiple*

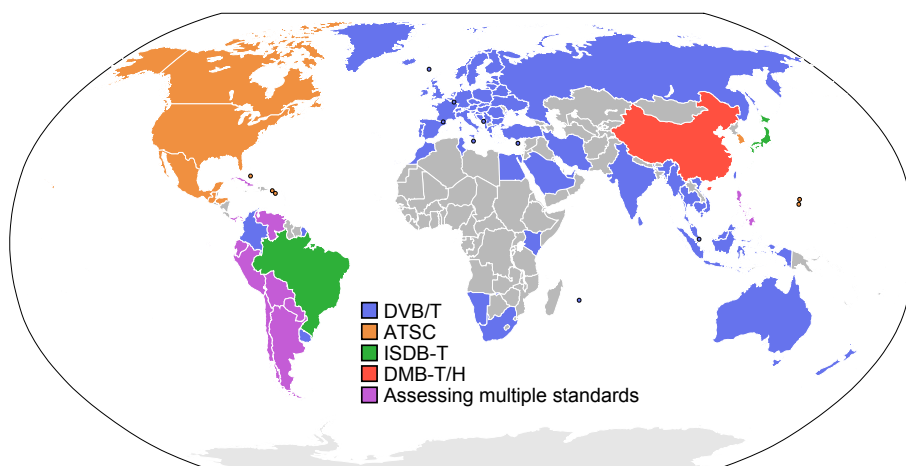


FIGURE 1.7 – Les différents standards de télévision numérique.

Sub-Nyquist Sampling Encoding system) au Japon, et la norme analogique HD-MAC (pour *High-Definition Multiplexed Analogue Components*) en Europe, notamment utilisée pour retransmettre les Jeux Olympiques de Barcelone en 1992. Ces systèmes analogiques n'ont cependant pas rencontrés le succès escompté, principalement à cause d'un coût élevé et d'un manque de contenu haute-définition disponible. Finalement, ces projets seront abandonnés au profit de la télévision numérique.

Depuis maintenant 10 ans, la télévision numérique se déploie sur la majeure partie du globe (figure 1.7) et les précédents formats établis depuis cinquante ans sont progressivement abandonnés. La fin du NTSC est prévue cette année 2009 aux États-Unis, et en 2011 au Japon et au Canada. Le PAL/SECAM sera abandonné en 2009 en Allemagne et en Scandinavie, en 2010 en Espagne, en 2011 en France, en 2012 en Italie et au Royaume-Uni. La plupart de ces nouveaux formats sont encore diffusés actuellement dans des formats standard (576i, 480i) mais des diffusions haute-définition sont déjà mises en place. Dans un futur plus ou moins proche selon les régions du globe, l'offre numérique devrait être proposée dans des formats haute-définition (720p et 1080i dans un premier temps, puis 1080p à terme).

Bien entendu, cette révolution télévisuelle sera inutile si elle n'est pas suivie d'une révolution technique au niveau des dispositifs d'affichage. Comme on l'a montré précédemment, les écrans CRT ont atteint leurs limites en termes de taille et de résolution, et des écrans plus grands sont nécessaires pour profiter des bénéfices apportés par la télévision haute-définition. C'est dans ce contexte que de nombreuses nouvelles technologies d'affichage ont fait leur apparition.

1.4.4 Nouveaux écrans

Pour succéder au monopole de l'écran à tube cathodique dans le domaine de la télévision, de nombreuses technologies d'affichage ont été développées. Deux technologies à écrans plats se démarquent nettement en termes de succès, les écrans à cristaux liquides (en anglais *Liquid Crystal Display*,

Caractéristiques	CRT	PDP	LCD
Tailles disponibles	10-36	37-70	7-65
Luminosité	•	•	••••
Contraste dans une pièce éclairé	•	•	•••••
Contraste dans une pièce sombre	•••	•••	•••••
Angle de vue	•••••	•••••	••••
Images en mouvement	•••••	•••••	•••
Images fixes	••	•	•••••
Durée de vie	•••••	•••	•••••
Poids / Encombrement	••	•••	•••••

TABLE 1.1 – Comparaison des différents types de téléviseurs [Display Search 05].

LCD) et les écrans plasma (en anglais *Plasma Display Panel*, PDP). On peut également ajouter les systèmes de projection et de rétro-projection qui utilisent principalement les technologies CRT, LCD, DLP (*Digital Light Processing*), ou LCoS (*Liquid Crystal on Silicon*). D'autres technologies prometteuses ont été annoncées ces dernières années, notamment les technologies SED (*Surface-conduction Electron-emitter Display*) et FED (*Field Emission Display*) toutes deux inspirées de la technologie CRT. Néanmoins il semblerait que les efforts de développement de ces deux techniques aient été stoppés face à la concurrence des écrans à cristaux liquides et plasma. Enfin, les écrans à diodes électro-luminescentes organiques (en anglais *Organic Light-Emitting Diode*, OLED) sont en train de faire leur apparition notamment pour les petits appareils (téléphones et ordinateurs portables principalement), ils devraient être introduits sur le marché des téléviseurs dans les années à venir.

Les qualités et défauts des trois types d'écrans utilisés en télévision sont récapitulés dans le tableau 1.1. Le principal défaut de la technologie plasma réside dans la résolution des écrans : pour des raisons techniques, il est difficile et coûteux de faire des écrans plasma haute définition de petite tailles. C'est pourquoi ce type d'écran n'est pas présent sur le marché des écrans d'ordinateurs (portables ou de bureau) où le LCD constitue la seule alternative aux écrans CRT. De la même façon, les téléviseurs plasma « Full-HD » (résolution native de 1920×1080) ne sont disponibles qu'à partir d'une taille d'environ 50" en diagonale. Les écrans LCD « Full-HD », de tailles moyennes (autour de 35") et donc moins chers, connaissent un large succès commercial au détriment des écrans plasma haut de gamme. En ce qui concerne la qualité d'affichage, le LCD est supérieur au plasma en termes de luminosité et de contraste, cependant il a quelques défauts notamment concernant la reproduction du mouvement et l'angle de vue.

Au fil des années, le LCD s'est progressivement installé sur le marché des téléviseurs jusqu'à devenir depuis fin 2007 le type de téléviseur le plus vendu au monde (figure 1.8). Moins d'un an plus tard, il représente plus de 80% des ventes de téléviseurs en Europe comme par exemple dans les îles britanniques (figure 1.9). En d'autres termes, le LCD est sorti grand vainqueur parmi la liste des prétendants au remplacement du CRT.

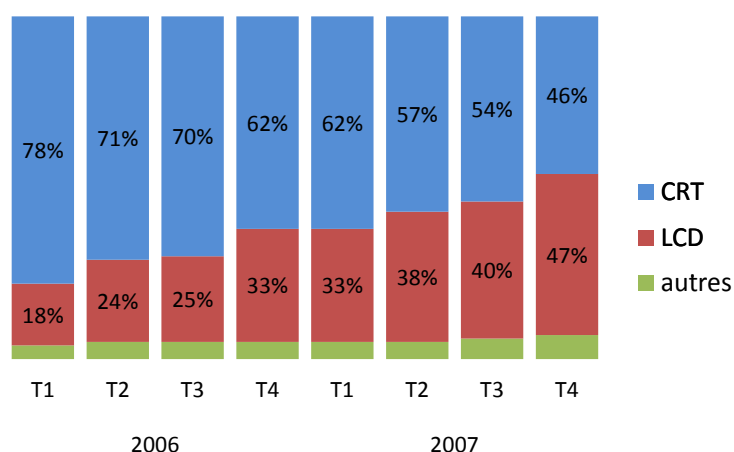


FIGURE 1.8 – Évolution par trimestres, en 2006 et 2007, du nombre de téléviseurs CRT et LCD vendus sur la planète, exprimées en pourcentage des ventes totales de téléviseurs [LCDTVA 08].

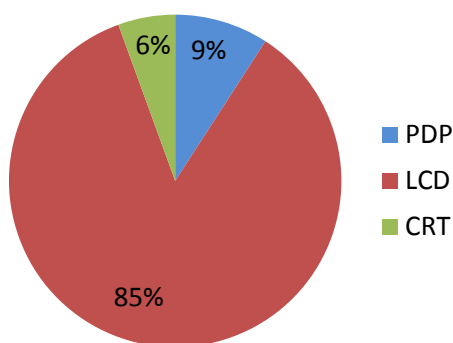


FIGURE 1.9 – Répartition par types d'écran des ventes de téléviseurs en Grande Bretagne et en Irlande pour le troisième trimestre 2008 [UKDisplay 09].

1.5 Conclusion

L'évolution de la télévision vers la haute-définition et les formats numériques, ainsi que la démocratisation des ordinateurs de bureau, ont rendu obsolètes les écrans à tube cathodique depuis la fin du 20^{ème} siècle. L'augmentation nécessaire de la résolution d'affichage et de la taille de l'écran, dans le but d'améliorer la qualité d'usage, a fait émerger de nouveaux écrans, plus légers, moins encombrants et plus économiques : les écrans plats. Parmi les nouveaux prétendants, l'écran à cristaux liquides est rapidement devenu le leader incontestable des écrans multimédia (téléviseurs, moniteurs de bureau, ordinateurs portables, etc.).

Malgré ses limites actuelles, l'écran CRT est le fruit d'un siècle de recherches et d'améliorations permanentes pour en faire un écran le plus adapté possible à l'observateur humain. À l'inverse, les écrans à cristaux liquides sont issus d'une technologie relativement jeune. Leur utilisation a longtemps été cantonnée à l'affichage de contenus fixes, et ils possèdent encore de nombreux défauts.

Ce nouveau système d'affichage électronique doit donc encore être amélioré et perfectionné pour atteindre une qualité digne de son prédécesseur. Pour identifier et résoudre les principaux défauts visuels des écrans à cristaux liquides, il est fondamental de comprendre le fonctionnement de ces écrans. C'est l'objet du second chapitre de ce mémoire.

Chapitre 2

Écrans à cristaux liquides, un nouveau contexte

Sommaire

2.1	Introduction	23
2.2	Les cristaux liquides	24
2.2.1	La découverte des cristaux liquides	24
2.2.2	Le cristal liquide en quelques mots	24
2.2.3	Application aux écrans	25
2.3	Fonctionnement d'un écran à cristaux liquides	25
2.3.1	Le panneau d'affichage	26
2.3.2	L'éclairage	29
2.3.3	L'adressage des pixels	34
2.4	Une nouvelle technologie à évaluer	35
2.4.1	De nouvelles conditions de visualisation	35
2.4.2	De nouvelles distorsions visuelles	37
2.4.3	Problématique	38
2.5	Conclusion	39

2.1 Introduction

Bien qu'ils furent découverts dès la fin du 19^{ème} siècle, il fallut attendre près d'un siècle, en 1968, pour que les cristaux liquides soient utilisés dans un dispositif d'affichage. À partir de cette date, les écrans à cristaux liquides, ou LCD pour *Liquid Crystal Display*, n'ont cessé de se développer. Leur faible consommation et leur encombrement modéré les ont tout naturellement désignés comme les afficheurs par excellence des appareils portables (montre, calculatrice, téléphone mobile, etc.). La démocratisation des ordinateurs personnels, puis des ordinateurs portables dans les années 1990, leur permet de pénétrer le marché des écrans multimédia. Avec l'arrivée de la télévision haute définition au début du 21^{ème} siècle, leur grande taille leur permet de s'imposer dans les salons en remplacement des téléviseurs à tube cathodique.

Dans ce chapitre, nous dressons un bref historique des cristaux liquides depuis leur découverte jusqu'à leur utilisation dans le domaine des écrans de télévision. Les écrans à cristaux liquides utilisés comme téléviseurs ou moniteurs informatiques sont ensuite décrits, notamment la structure du pixel, l'éclairage et l'adressage matriciel actif. Enfin, ces considérations techniques nous amèneront à observer les principales différences entre la technologie émergente des cristaux liquides et celle des écrans à tubes cathodiques. Nous identifierons alors les imperfections visuelles introduites par les écrans LCD.

2.2 Les cristaux liquides

2.2.1 La découverte des cristaux liquides

Dès la seconde moitié du 19^{ème} siècle, des travaux scientifiques en chimie et en biologie font état du comportement étrange de certains matériaux autour de leur température de fusion. En 1850, W. Heintz observe que le passage de l'état solide à l'état liquide de la stéarine se fait via un état liquide, tantôt trouble, tantôt opaque, selon la température. D'autres études rapportent également les observations de caractéristiques optiques anisotropes dans certains matériaux organiques « liquides », un comportement qu'on pensait alors réservé aux phases cristallines.

En 1888, alors qu'il étudie le cholestérol dans les végétaux, Friedrich Reinitzer, un botaniste autrichien découvre une molécule qui possède deux points de fusion [Reinitzer 88]. Pour tenter d'expliquer ce phénomène il collabore avec le physicien allemand Otto Lehmann. Leurs travaux mènent finalement à la découverte des cristaux liquides et à la compréhension fondamentale de ce nouvel état de la matière. Après cette découverte, la recherche sur les cristaux liquides se développe rapidement pour finalement conduire à la classification des différentes phases par Georges Friedel en 1922 [Friedel 22].

2.2.2 Le cristal liquide en quelques mots

Les molécules qui compose un solide sont positionnée et orientée de manière ordonnée. Inversement, dans un liquide les molécules peuvent se déplacer librement de façon aléatoire. Entre ces deux états de la matière, il existe des phases condensées qui présentent des caractéristiques intermédiaires, comme par exemple le cristal liquide. On distingue différents types de cristaux liquides selon la forme des molécules, leur agencement, leur orientation directrice, etc.

Propriétés électro-magnétiques La susceptibilité électro-magnétique des cristaux liquides n'est pas la même suivant l'axe parallèle et suivant l'axe perpendiculaire à l'axe directeur des molécules. Cette anisotropie permet de déterminer la réorientation finale des molécules soumises à un champ électro-magnétique. Si la susceptibilité le long de l'axe est plus grande que la susceptibilité normale à l'axe alors les molécules s'aligneront parallèlement au champ et vice-versa.

Propriétés optiques Le cristal liquide est un milieu biréfringent (l'indice de réfraction du milieu n'est pas unique). Une lumière polarisée parallèlement à l'axe directeur du cristal est soumise à

un indice de réfraction appelé l'indice de réfraction extraordinaire. Une lumière polarisée perpendiculairement à l'axe directeur est soumise à un autre indice, l'indice de réfraction ordinaire. Par conséquent, une lumière polarisée parallèlement ou perpendiculairement à l'axe directeur ne verra pas son état de polarisation changer. Cependant, si l'angle de polarisation de la lumière incidente est différent de 0° ou 90° , alors un retard de phase sera observé à la sortie et la lumière deviendra polarisée elliptiquement.

Ce sont ces propriétés électro-optiques qui sont utilisées dans les écrans LCD.

2.2.3 Application aux écrans

Le 20^{ème} siècle fut le siècle de l'écran CRT dans le domaine de la télévision (chapitre 1). Cependant, malgré ses avantages évidents (en terme de qualité d'image, de reproduction du mouvement, d'angle de vue, etc.), sa forte consommation électrique et sa forme encombrante ont contraint les fabricants d'écrans à trouver d'autres alternatives, à l'heure où la haute-définition se déploie.

Malgré la découverte relativement ancienne des cristaux liquides, leur potentiel dans le domaine de l'affichage n'a été réalisée qu'en 1968. Cette année là, George Heilmeyer présenta le premier afficheur à cristaux liquides basé sur l'effet de dispersion dynamique qu'il venait d'inventer. En 1970, le premier écran à cristaux liquides est développé par Optel Corporation. Dans les années 1970, les Japonais entrent sur le marché des écrans à cristaux liquides, Semko-Epson l'utilise sur les montres numériques (1972), Casio et Sharp sur les calculatrices (1973). En 1971, Schadt et Helfrich d'un côté et Fergason de l'autre inventent, indépendamment et presque au même moment, le mode nématique torsadé (en anglais *twisted nematic*, TN) encore utilisé dans les écrans d'aujourd'hui.

Au fur et à mesure que la technologie devient de plus en plus mature, l'utilisation des écrans LCD est en augmentation constante, non seulement dans les domaines conventionnels (informatique, télévision, etc.), mais également dans des domaines plus pointus tels que l'imagerie médicale ou la 3-D.

2.3 Fonctionnement d'un écran à cristaux liquides

Les deux principaux éléments d'un système d'affichage à cristaux liquide sont les suivants :

- Un ensemble de cellules de cristaux liquides, le panneau d'affichage, qui constituent les éléments de base de l'image affichée, et dont la transmittance est contrôlée de manière à laisser passer (cellule éclairée) ou à stopper (cellule éteinte) la lumière.
- Une source lumineuse, qui permet de révéler l'image affichée par les cellules de cristaux liquides ; elle peut-être incorporée à l'afficheur ou bien être externe (luminosité ambiante).

L'écran dans son ensemble est constitué de la juxtaposition en sandwich de plusieurs surfaces planes (cf. 2.1), dont le panneau d'affichage constitué des cellules de cristaux liquides. Devant le panneau, divers films permettent l'amélioration de la qualité d'affichage (anti-reflet, anti-éblouissement, augmentation du rendement lumineux, etc.). Derrière le panneau, la source lumineuse est diffusée de manière à obtenir un éclairage uniforme sur toute la surface de l'écran. Chaque couche est extrêmement fine, l'épaisseur de l'écran est due principalement au dispositif d'éclairage. Dans la suite de

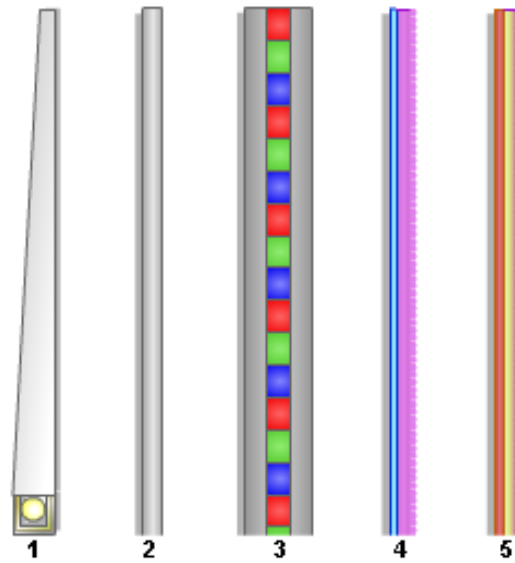


FIGURE 2.1 – Structure d'un écran à cristaux liquides (coupe latérale). L'écran est constitué d'un dispositif d'éclairage (1), d'une surface diffusante (2), du panneau de cellules de cristaux liquides (3), de plusieurs films améliorant l'affichage (4) et d'une surface protectrice (5). (tiré de [Vikuiti 05])

cette section, nous allons décrire plus précisément ces éléments en nous focalisant sur les écrans à cristaux liquides utilisés comme téléviseurs ou comme moniteurs informatiques.

2.3.1 Le panneau d'affichage

Un écran à cristaux liquides affiche des images par l'intermédiaire du panneau d'affichage. Celui-ci est constitué d'une matrice de pixels, généralement disposés orthogonalement. Un pixel représente un échantillon ponctuel de l'image affichée, défini par sa position et sa valeur (sa couleur). Chaque pixel est constitué de trois cellules de cristaux liquides, qu'on désigne généralement par le terme « sous-pixel ». À chaque sous-pixel est associé un filtre chromatique, de couleur rouge, verte ou bleue. C'est le mélange de ces trois couleurs qui permet de former la couleur finale du pixel (cf. section 1.2.2.1 et figure 1.2).

Les sous-pixels, ou plutôt les cellules de cristaux liquides qui les composent, ne sont pas des sources lumineuses à proprement parler, mais doivent être vus comme des « valves lumineuses », qui permettent le passage d'un flux lumineux plus ou moins important. Ce comportement est différent de celui des écrans CRT pour lesquels les pixels sont constitués de luminophores phosphorescents qui émettent le flux lumineux.

2.3.1.1 La structure d'un sous-pixel

Si l'on suit le trajet de la lumière depuis la source lumineuse jusqu'à l'extérieur de l'écran, les éléments du sous-pixel successivement traversés par le flux lumineux sont les suivants : un polariseur d'entrée, une électrode transparente, la cellule de cristaux liquides, une électrode transparente et un

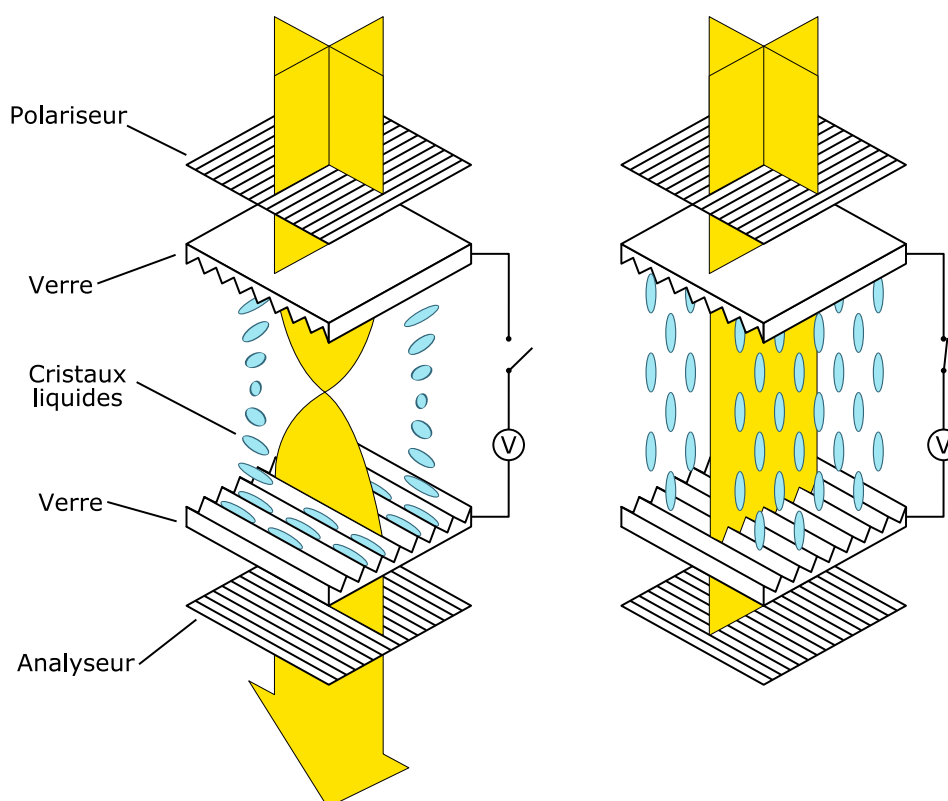


FIGURE 2.2 – Illustration du fonctionnement d'un sous-pixel dans un écran à cristaux liquides dans le cas de cristaux nématiques torsadés (en anglais *twisted nematic*, TN).

polariseur de sortie (ou analyseur). Plusieurs couches d'un substrat en verre sont intercalées entre ces différents éléments.

La figure 2.2 illustre le fonctionnement d'un sous-pixel dans le cas d'un cristal liquide de type nématique torsadé. La lumière incidente est polarisée avant de pénétrer dans la cellule de cristaux liquides. Au repos, les molécules de cristaux liquides ont une organisation hélicoïdale qui va modifier la polarité de la lumière lui permettant ainsi de passer à travers le polariseur de sortie (l'analyseur) : le pixel est allumé. Si on applique une tension entre les deux électrodes transparentes, les molécules de cristaux liquides s'alignent et la polarité de la lumière n'est plus modifiée. La lumière est alors bloquée par l'analyseur : le pixel est éteint. L'intensité lumineuse émise par le pixel peut être réglée en variant la tension aux bornes des électrodes.

2.3.1.2 Les différentes technologies de cristaux liquides

On distingue différents types de cristaux liquides en fonction de la forme des molécules, de leur agencement, de leur orientation directrice, etc. Certaines caractéristiques essentielles de l'écran sont grandement influencées par le choix de la technologie, notamment :

- le temps de réponse du pixel est directement lié au temps nécessaire aux molécules de cristaux liquides pour s'aligner lorsqu'elles sont soumises à un champ électro-magnétique ;
- l'angle de vue, c'est à dire l'angle sous lequel l'image affichée peut-être visualisée de manière acceptable, dépend lui aussi du type de cristaux et de leur orientation ;
- l'intensité du noir, et par conséquent le contraste¹, dépend de la transmittance du pixel lorsqu'il est opaque (car le rétro-éclairage reste allumé en permanence d'où la nécessité d'avoir une transmittance la plus faible possible en état d'opacité), cette transmittance est différente selon la structure du cristal liquide choisi ;
- la luminosité maximale de l'écran, limitée elle-aussi par la transmittance du pixel dans son état transparent.

On distingue principalement trois grandes familles parmi les technologie à cristaux liquides utilisées dans les téléviseurs et moniteurs à cristaux liquides [Gu 06]. Nous n'allons pas ici entrer dans les détails structurels de chaque technologie mais simplement comparer leurs performances. Le tableau 2.1 résume brièvement les principales qualités et les principaux défauts de ces différents modes en terme de reproduction d'image.

Twisted nematic (TS) C'est la technologie la plus courante et la plus ancienne. Elle utilise les cristaux liquides du même nom dont le fonctionnement assez simple est connu depuis des années (cf. figure 2.2). Cette technologie bénéficie d'un très bon temps de réponse et d'une bonne luminosité mais souffre d'un angle de vue très limité et d'un noir peu profond. Pour ces raisons son utilisation est principalement limitée aux moniteurs de bureau. Il convient également de noter que le pixel au repos est « allumé », il faut lui appliquer un champ électrique pour l'éteindre.

In-plane switching (IPS) Cette technologie utilise des cristaux liquides qui ont la particularité de changer de transmittance tout en restant alignés dans un même plan. Cela permet d'obtenir un très bon angle de vue, une caractéristique très importante en télévision. En contrepartie, le temps de réponse de ces écrans est assez faible. Un des gros inconvénients du mode IPS est qu'il nécessite la présence de deux transistors par pixel, réduisant d'autant la surface utile de ce dernier. La puissance utile de l'éclairage s'en trouve donc réduite d'autant. Un certain nombre de modifications ont été apportées à ce mode par la suite, notamment avec le Super-IPS (S-IPS) qui améliore le procédé de manière à obtenir une plus grande luminosité, tout en améliorant la réponse grâce à des techniques de compensation du temps de réponse.

Vertical-Alignment (VA) On retrouve la technologie VA dans les téléviseurs principalement. Ce mode possède les mêmes avantages que le mode IPS en terme d'angle de vision et de niveau du noir. De plus, la structure du cristal, et notamment son alignement vertical au repos, permet d'obtenir des temps de réponse inférieurs à ceux du mode IPS. Ici encore, de nombreuses évolutions existent. On peut citer notamment le mode *Multi-domain Vertical Alignment* (MVA) qui a longtemps été annoncé comme le successeur du mode TN mais en vain, le mode *Patterned Vertical Alignment* (PVA) assez

1. Le contraste d'un écran, dans son sens général, correspond au rapport de la luminance du blanc sur la luminance du noir.

	État au repos	Temps de réponse	Angle	Luminosité	Noir	Utilisation p ^{ale}
TN	blanc	++	-	++	-	moniteurs
IPS	noir	- -	++	-	+	moniteurs, TV
S-IPS	noir	-	+++	+	+	moniteurs, TV
MVA	noir	+	++	-	+	moniteurs, TV
PVA	noir	+	++	+	++	TV
ASV	noir	+	+++	++	++	TV

TABLE 2.1 – Comparaison des différents modes de cristaux liquides utilisés dans les écrans LCD.

similaire au mode MVA mais avec un meilleur contraste et le mode ASV (pour *Advanced Super View* et surtout *Axially Symmetric Vertical alignment*) qui permet d'atteindre des contrastes très élevés tout en préservant un angle de vision quasiment complet.

2.3.2 L'éclairage

Lorsqu'on s'intéresse au dispositif d'éclairage d'un afficheur LCD, on distingue généralement trois types d'écran à cristaux liquides. Ces différentes configurations sont représentées figure 2.3. Dans les écrans transmissifs, la lumière est issue d'une source lumineuse intégrée à l'écran et située à l'arrière du panneau d'affichage. Ce type d'éclairage offre la plus grande luminance, en particulier pour les écrans de grande taille, il est utilisé dans les moniteurs d'ordinateur et les téléviseurs LCD. Différents types de sources lumineuses existent parmi lesquelles les tubes fluorescents et les diode électro-luminescentes. Il existe également des écrans réflectifs, à l'intérieur desquels un miroir réfléchit la lumière provenant de l'extérieur à travers le panneau d'affichage. Cette approche est généralement utilisée pour de petits écrans monochromes montés sur des appareils qui ont de faibles réserves d'énergie, par exemple les montres ou les calculatrices. Enfin les écrans transflectifs sont une combinaison des deux techniques précédentes. Leur structure est semblable à celle des écrans transmissifs mais avec un transflecteur derrière le panneau de cristaux liquides. Le transflecteur est généralement un semi-miroir qui permet la réflexion de la lumière ambiante tout en laissant passer la lumière issue du rétro-éclairage. L'utilisateur peut généralement choisir d'utiliser ou non le rétro-éclairage selon qu'il se trouve dans un environnement plus ou moins éclairé. On trouve ce type d'écran à économie d'énergie sur les petits appareils à batterie tels que les téléphones mobiles ou les lecteurs multimédia.

2.3.2.1 Le rétro-éclairage

Dans le cadre de notre étude, nous nous intéressons aux écrans transmissifs, tels que les moniteurs de bureau, les écrans d'ordinateur portables et les téléviseurs LCD. Parce qu'il est situé à l'arrière du panneau d'affichage, le dispositif d'éclairage des écrans transmissifs est appelé rétro-éclairage. Il

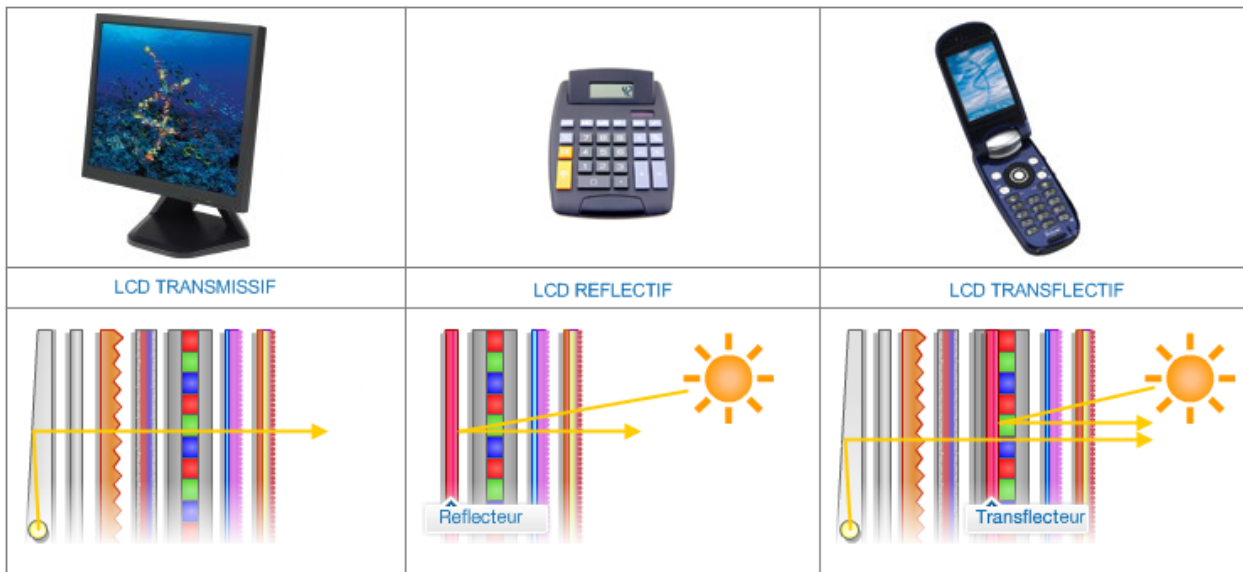


FIGURE 2.3 – Les trois différentes configurations d'éclairage sur LCD. (tiré de [Vikuiti 05])

peut être direct ou latéral. Dans le premier cas les sources lumineuses se situent directement derrière le panneau, dans le second cas elles sont installées sur les cotés de l'écran ou sous celui-ci.

La figure 2.4 illustre les configurations les plus répandues. Sur les moniteurs de bureau (a), la lumière est émise par plusieurs sources latérales situées de chaque coté de l'écran. Les rayons lumineux sont maintenus dans un guide de lumière duquel il ne peuvent s'échapper qu'à partir d'un certain angle incident, dans le cas contraire la réflexion est totale et les rayons restent dans le guide. Lorsque la lumière atteint un des points d'extraction situés sur la face arrière du guide, elle est diffusée vers le panneau d'affichage. Une plaque de diffusion permet d'uniformiser l'illumination sur toute la surface de l'écran.

L'architecture est similaire (b) pour les écrans d'appareils portatifs (ordinateur, téléphone, etc.) si ce n'est que la source lumineuse est généralement située sous l'écran. Pour minimiser l'épaisseur de l'écran, on utilise un guide de lumière de forme biseautée sans points d'extraction. A chaque réflexion, l'angle d'incidence des rayons lumineux diminue jusqu'à permettre l'échappement. L'angle de réfraction étant assez élevé (typiquement autour de 70°), un film prismatique est ajouté de manière à dévier les rayons en direction de l'utilisateur.

Pour les téléviseurs enfin (c), une forte luminance est généralement nécessaire, les sources lumineuses sont alors situées directement derrière le panneau dans une cage réfléchissante. Ici, le processus de diffusion doit être très important de manière à rendre invisible les sources lumineuses situées directement dans le champ visuel de l'observateur. Néanmoins, des problèmes d'uniformité peuvent apparaître, notamment lorsque les différentes sources lumineuses utilisées vieillissent à des vitesses différentes.

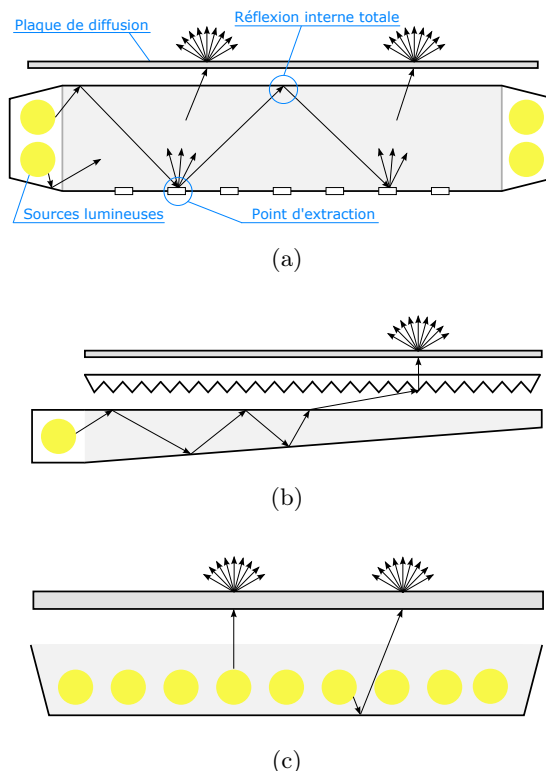


FIGURE 2.4 – Différents types de rétro-éclairage : *latéral* sur les moniteurs de bureau (a) ou sur les ordinateurs portables (b), *direct* sur les téléviseurs (c). (inspirée de [Vikuiti 05])

2.3.2.2 Les différents types de sources lumineuses

Le choix de la source lumineuse utilisée dans le rétro-éclairage des écrans à cristaux liquides dépend de différents paramètres parmi lesquels le rendu visuel, le coût, la consommation électrique. Dans le cas des écrans d'ordinateur ou des téléviseurs on rencontre principalement deux types de sources :

- les tubes fluorescents à cathode froide (TF CF)
- les diodes électro-luminescentes (DEL)

Les tubes fluorescents à cathode froide sont les plus répandus. Ils sont fins et légers et émettent une lumière blanche très lumineuse. Leur consommation en énergie est assez faible et leur durée de vie moyenne est de plusieurs dizaines de milliers d'heures. Cependant, leur luminance est très sensible aux faibles températures et ils nécessitent une alimentation haute-tension assez complexe.

Autrefois réservées aux petits écrans qui ne nécessitaient pas de fortes luminances (montres, calculatrices), les diodes électro-luminescentes sont de plus en plus utilisées. Les avancées récentes dans le domaine des DEL (notamment le développement des DEL de couleur bleue et des DEL blanches) a permis de les utiliser sur des petits écrans (téléphones mobiles, ordinateurs portables) et elles devraient équiper les moniteurs et téléviseurs haut de gamme dans les années à venir. Leur longue durée de vie (50000 heures minimum) et leur alimentation en faible tension jouent également en leur faveur. Néanmoins, la génération actuelle d'éclairage par DEL est assez gourmande en énergie bien qu'il soit théoriquement possible d'optimiser la consommation électrique pour atteindre des niveaux inférieurs à ceux des TF CF.

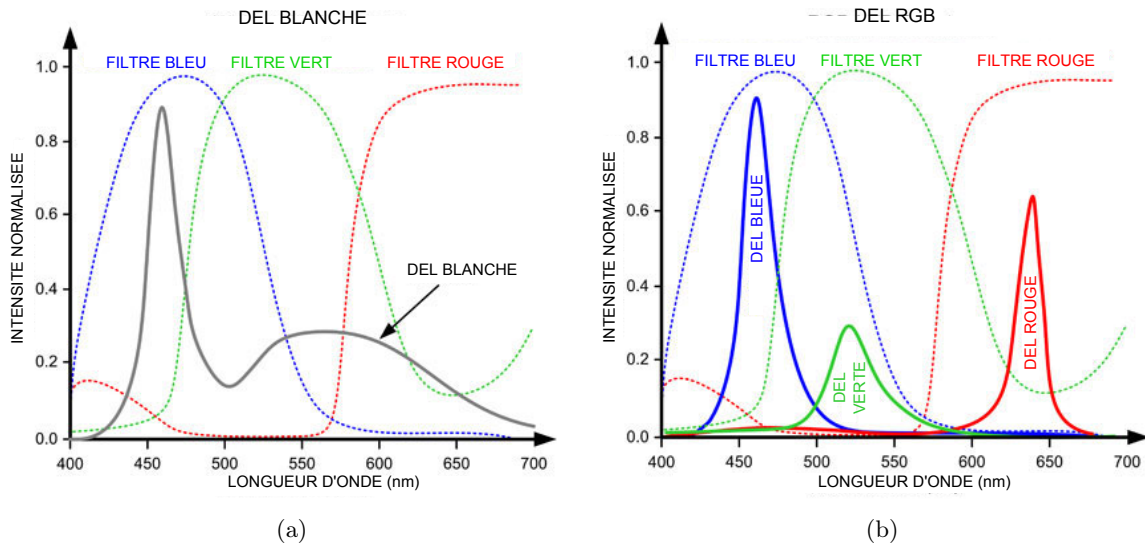


FIGURE 2.5 – Densité spectrale de puissance d’un éclairage à DEL blanches (a) et d’un éclairage utilisant un mélange de DEL rouges, verts et bleues (b). La transmittance des filtres colorés correspondant à chaque sous-pixel est également tracée (en pointillés).

La principale différence entre le TFCF et la DEL intervient aux niveaux des couleurs affichées par l’écran. Les TFCF émettent une lumière blanche qui couvre l’ensemble du spectre visible. Chaque composante chromatique (rouge, vert et bleu) est créée par l’intermédiaire du filtre chromatique dont est équipé chaque sous-pixel. Étant donné le large spectre couvert par la source blanche, les composantes chromatiques ne sont pas des couleurs pures, l’ensemble des couleurs reproductibles par l’écran est donc assez réduit. Dans le cas d’un écran utilisant un rétro-éclairage à DEL blanches le problème est sensiblement le même comme le montre la figure 2.5a. La solution utilisée pour augmenter la gamme de couleurs de l’écran est donc l’utilisation comme source lumineuse d’un mélange de DEL rouge, vert et bleu. La densité spectrale de puissance de l’éclairage complet est alors composée de trois pics relativement étroits et les composantes chromatiques sont par conséquent relativement pures (figure 2.5b). La gamme des couleurs reproductibles par l’écran est alors bien plus étendue que dans le cas d’une lumière blanche. La figure 2.6 compare les gammes de couleurs reproductibles pour chaque type d’éclairage.

Un autre avantage d’un rétro-éclairage à DEL est la possibilité, dans le cas d’un rétro-éclairage direct, d’adapter localement la luminance. La figure 2.7a illustre la structure d’un rétro-éclairage direct constitué de diodes électro-luminescentes. La luminance de chaque DEL peut être modulée de manière à générer une image de luminance basse résolution du contenu [Trentacoste 07] (cf. figure 2.7b). Ce dispositif permet d’augmenter le contraste de l’image en réduisant le niveau de luminosité des zones sombres et d’afficher ainsi des images à grande dynamique (cf. figure 2.7b).

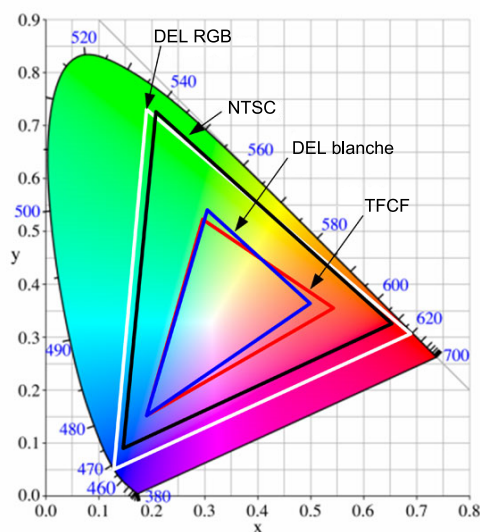


FIGURE 2.6 – Représentation des couleurs dans l'espace de coordonnées (x,y) . La zone colorée représente l'ensemble des couleurs perceptibles et le triangle noir l'ensemble des couleurs potentiellement présentes dans le signal vidéo (format NTSC). L'ensemble des couleurs reproductibles par un écran équipé d'un éclairage blanc (TFCF ou DEL blanches) est assez réduit tandis que celui d'un écran équipé d'un mélange de DEL colorées (DEL RGB) est capable d'afficher toutes les couleurs du format NTSC voire même au delà.

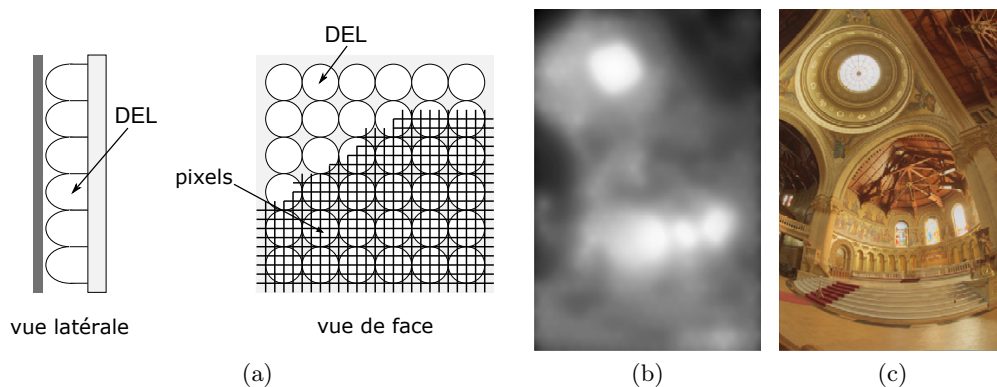


FIGURE 2.7 – Rétro-éclairage à DEL constituant un dispositif imageur basse résolution (a) ; exemple d'image basse résolution générée par le rétro-éclairage à DEL (b) pour afficher une image à grande gamme dynamique (c).

2.3.3 L'adressage des pixels

2.3.3.1 Les modes d'adressage

Le mode d'adressage d'un écran correspond à la méthode utilisée pour affecter leur valeur à l'ensemble des pixels d'un écran. C'est une caractéristique importante de l'écran puisqu'elle va déterminer la façon dont est formée l'image. On distingue généralement trois modes d'affichage : l'adressage direct, l'adressage raster, et enfin l'adressage matriciel.

L'adressage direct consiste à contrôler individuellement chaque pixel de l'écran. La valeur du pixel est donc maintenue, et mise à jour de façon continue. Pour un écran de $m \times n$ pixels, ce mode d'adressage nécessite $m \times n$ signaux de contrôle et autant de connexions. Il est donc principalement utilisé pour les afficheurs avec un petit nombre de segments comme les montres et les calculatrices. Pour les écrans composés d'un nombre important de pixels, il est inconcevable de pouvoir utiliser ce mode d'adressage.

L'adressage raster est le mode d'adressage utilisé dans les écrans CRT. Les pixels sont alors parcourus séquentiellement un par un (balayage), et le signal d'intensité est modulé au fur et à mesure pour activer le pixel courant. La persistance temporelle du pixel (les luminophores du CRT par exemple) associée à la persistance rétinienne permet le maintien de l'état du pixel entre deux balayages. Seulement trois signaux de contrôle sont nécessaires pour ce mode d'adressage : le signal de balayage horizontal, le signal de repositionnement vertical et le signal d'intensité. Le timing entre ces trois signaux est très important.

L'adressage matriciel utilise un signal de contrôle pour chaque ligne et un signal de contrôle pour chaque colonne. À chaque pixel correspond la combinaison d'un signal de ligne et d'un signal de colonne. Un écran de $m \times n$ pixels nécessite donc $m + n$ signaux de contrôle. Comme pour l'affichage raster, le signal de contrôle de chaque pixel n'est pas permanent. Néanmoins, l'état du pixel peut être maintenu de différentes façons. Dans le cas de l'adressage matriciel actif, une sorte de condensateur est utilisé pour maintenir le niveau du signal. Dans le cas où le pixel ne peut prendre que deux valeurs (allumé ou éteint) on peut également utiliser des éléments bi-stables, on parle alors d'adressage matriciel passif.

2.3.3.2 La matrice active

Les écrans à cristaux liquides monochromes des premiers ordinateurs portables utilisaient un adressage matriciel passif. Les défauts de ce mode d'adressage étaient nombreux : faible contraste, échelle de gris réduite, etc. L'utilisation de l'adressage matriciel actif a révolutionné les écrans à cristaux

liquides et a permis leur usage en tant que téléviseurs et moniteurs d'ordinateur. Les écrans modernes, utilisant ce mode d'adressage, sont appelés LCD à matrice active (en anglais *active-matrix liquid crystal display*, AM-LCD).

La généralisation des écrans à matrice active a été rendu possible grâce à l'utilisation de transistors sur couche mince (en anglais, *thin-film transistor*, TFT) on parle alors de TFT-LCD, cette expression étant devenue équivalente à AM-LCD bien qu'un LCD à matrice active puisse théoriquement intégrer d'autres composants (des diodes par exemple). Mais en pratique tous les AM-LCD modernes sont des TFT-LCD.

Le principe des TFT-LCD consiste donc à associer un transistor à chaque pixel, pour servir d'interrupteur et ainsi isoler le pixel du reste du circuit. Ces transistors sont fabriqués à partir d'une couche mince de silicium déposée sur le substrat en verre. Le reste du pixel est constituée de la cellule de cristaux liquides (cf. figure 2.2) qui joue le rôle de condensateur, et qui est prise en sandwich entre les deux électrodes transparentes. Grâce à leur fabrication en couches minces, les transistors n'occupent qu'une petite partie de la surface d'un pixel, et le reste du silicium est enlevé pour permettre à la lumière de passer. Néanmoins, leur présence réduit quelque peu la luminosité de l'écran.

2.4 Une nouvelle technologie à évaluer

Avec le déploiement des contenus haute-définition (TVHD, blu-ray, jeux vidéo, etc.), l'augmentation de la taille de l'écran est devenue nécessaire pour apprécier les avantages de l'augmentation de la résolution des images. Dans ces conditions, ses qualités en terme de design, d'encombrement et de poids ont permis à l'écran à cristaux liquides de conquérir le marché des téléviseurs, au détriment des anciens téléviseurs à tube cathodique. Néanmoins, son fonctionnement est très éloigné de celui de l'écran CRT. Depuis la structure individuelle d'un pixel, jusqu'à la méthode d'adressage, en passant par la source lumineuse, il existe de nombreuses différences entre les deux écrans.

La technologie d'affichage à cristaux liquides reste encore très récente comparée à la technologie CRT. Elle n'est pas directement adaptée à l'affichage de contenus télévisuels et de nombreuses distorsions visuelles restent encore à corriger. De plus, l'arrivée de la télévision haute-définition a changé les conditions de visualisation des téléspectateurs et cet aspect doit également être pris en compte pour l'évaluation de ces nouvelles distorsions visuelles.

2.4.1 De nouvelles conditions de visualisation

Le succès des écrans à cristaux liquides ces dernières années est en partie dû à l'arrivée de la télévision haute-définition. Dans ce contexte, de nouvelles conditions de visualisation sont introduites.

Sollicitation de la vision périphérique Pour que l'utilisateur final puisse éprouver des sensations proche du cinéma, l'augmentation de la résolution des images doit être associée à une augmentation de la taille de ces images dans son champ visuel. La figure 2.8 illustre ce changement de configuration. Grâce à l'augmentation de la résolution, l'observateur peut se placer plus près de

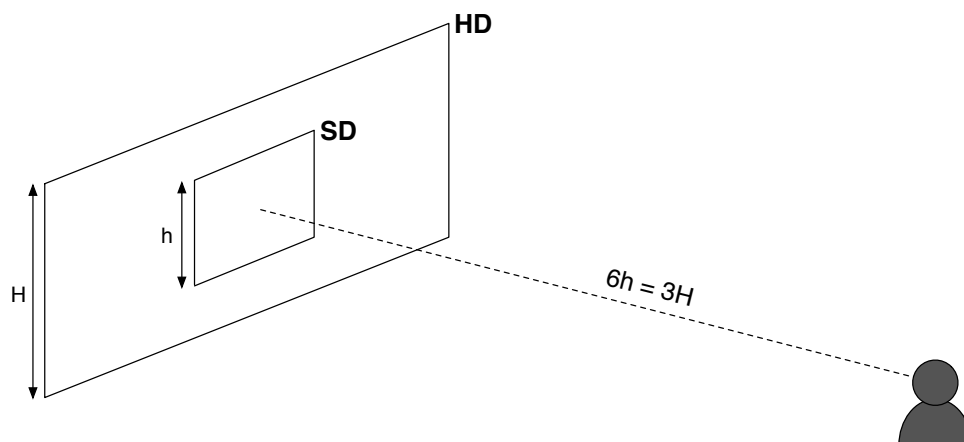


FIGURE 2.8 – Illustration du changement de configuration en télévision haute définition (TVHD) par rapport à la télévision standard (TVSD) en terme d’occupation du champ visuel. En TVSD, l’image est située à une distance égale à 6 fois la taille de l’écran, elle a une taille d’environ 10 degrés visuel. En TVHD, la distance de visualisation est réduite à 3 fois la taille de l’écran, l’image mesure alors 18 degrés visuels.

l’écran sans percevoir la structure des pixels de l’écran (cf. section 1.4). Ainsi en télévision haute-définition la distance optimale est de l’ordre de 3 à 4 fois la hauteur de l’image (selon les formats de diffusion) [ITU 98], la hauteur de l’image est alors de 18 degrés visuels. Par comparaison, la distance recommandée en télévision standard était de l’ordre de 6 fois la hauteur de l’image (l’image sous-tend alors un angle de l’ordre de 10 degrés visuels) [ITU 04].

Ce changement de configuration a un impact direct sur les conditions de visualisation de l’image affichée. Ce ne sont plus seulement les zones fovéale et péri-fovéale qui sont excitées mais également la zone para-fovéale de la rétine. Or la vision périphérique n’a pas les mêmes caractéristiques que la vision fovéale ou péri-fovéale, à cause la répartition inhomogène des photo-récepteurs (cf. annexe, section A.4).

Augmentation du niveau de luminosité Lorsqu’on regarde la télévision sur un écran à tube cathodique, la luminance maximale est limitée pour deux raisons. La première raison est technique, l’image est affichée pixel par pixel, et bien que la luminance émise par chaque pixel soit très intense, cette source de forte luminosité est limitée spatialement et temporellement. Par conséquent, la luminance moyenne sur l’ensemble de l’image est fortement réduite. La seconde raison est liée encore une fois au système visuel humain. La sensibilité au papillotement de l’œil est fortement dépendante de la luminosité (cf. section A.4.2.2). La luminance maximale d’un écran CRT utilisé en télévision est donc normalisée à 70 cd/m^2 [ITU 94, ITU 04], pour s’assurer que le papillotement soit imperceptible pour la grande majorité des observateurs.

Sur un écran à cristaux liquides, la luminance maximale de l’écran dépend exclusivement du rétro-éclairage utilisé (cf. section 2.3.2). Que ce soit dans le cas des TFCE ou dans le cas de DEL, la luminance maximale peut atteindre plusieurs centaines de cd/m^2 , voire même plus de 1000 cd/m^2

	CRT	LCD
Technologie	mature	récente
Adressage	raster	matriciel actif
Réponse temporelle	instantanée	lente
Papillotement	présent	aucun
Contraste	indépendant de l'angle de vue	dépendant de l'angle de vue
Noir parfait	possible	impossible

TABLE 2.2 – Différences visuelles principales entre CRT et LCD [Carrein 03].

sur les écrans les plus récents. De plus, grâce à l'adressage matriciel actif l'image est maintenue à l'écran entre deux rafraîchissements, il n'y a donc plus de risques de papillotement. Cependant, la réponse du système visuel humain est fortement dépendante de la luminance du stimulus (cf. annexe A). Par exemple, la loi de Weber-Fechner (cf. section A.5.1) ne s'applique plus pour des luminances supérieures à 1000 cd/m^2 , de même les fonctions de sensibilité au contraste sont différentes (cf. section A.5.2). Par conséquent, un certain nombre de propriétés du système visuel humain utilisées dans le domaine de télévision (notamment pour la quantification du signal télévisuel dans le schéma de codage) sont modifiées ou ne sont plus vérifiées sur un écran à cristaux liquides.

2.4.2 De nouvelles distorsions visuelles

Plusieurs études ont porté sur la comparaison entre la qualité d'affichage sur CRT et la qualité d'affichage sur LCD [Carrein 03, EBU 04, ITU 05]. Le tableau 2.2 recense les principales différences, vis à vis de la qualité perçue, entre ces deux types d'écrans. On peut notamment s'attarder sur les travaux d'un groupe de travail de l'Union Internationale des Télécommunications (ITU) [ITU 05] qui a mené une série de tests subjectifs avec un panel de 36 observateurs experts du domaine de la diffusion. Les résultats de cette évaluation subjective ont montré que, dans l'ensemble, la majorité des observateurs jugeait la qualité des images affichées sur LCD inférieure à celle des images affichées sur CRT. De nombreux défauts ont été recensés par les observateurs :

- Le flou de mouvement, qui est dû au maintien de l'image sur l'écran durant toute la période d'affichage, est le défaut le plus gênant. Lorsque l'observateur suit un objet en mouvement, ses yeux se déplacent sur l'écran de manière continue. À l'inverse, l'objet se déplace de manière discrète. Par conséquent, entre deux images, les yeux, en phase de suivi de mouvement, se déplacent sur un contenu momentanément fixe, qui devient flou sur la rétine.
- L'adressage matriciel impose l'affichage d'une image progressive, les contenus entrelacés doivent donc être désentrelacés par l'écran. Ce traitement n'est pas simple et doit être effectué en temps réel par l'écran. Pour les séquences avec des mouvements de fortes amplitudes, des distorsions gênantes peuvent apparaître.
- La restitution des couleurs est différente de celle des écrans CRT. En effet, les trois couleurs primaires d'un écran à cristaux liquides sont formées, dans la plupart des cas, par une lumière blanche qui passe à travers trois filtres chromatiques (cf. section 2.3.2). Les densités spectrales de

- puissances des trois sources chromatiques ainsi obtenues sont très différentes des densités spectrales de puissance des luminophores phosphorescents présents sur les écrans à tube cathodique.
- La profondeur du noir est très insuffisante sur les écrans à cristaux liquides. En effet, pour éteindre un pixel il faut régler sa transmittance à zéro, mais le rétro-éclairage reste allumé et une légère luminosité continue d'être perçue. À l'inverse, le noir d'un écran CRT peut correspondre à une absence totale de lumière si l'utilisateur le souhaite. Ce problème peut également se répercuter sur l'équilibrage des couleurs, particulièrement pour de faibles luminances.
 - Enfin des problèmes d'uniformité peuvent apparaître, ils sont de deux types : un mauvais réglage des TFCF constituant le rétro-éclairage peut conduire à une non-uniformité de la luminance sur l'ensemble de l'écran ; et l'hétérogénéité de la densité des cristaux liquides peut conduire à des différences de transmittance d'une région de l'écran à l'autre.

2.4.3 Problématique

Les images et les séquences vidéo affichées par un écran doivent être présentées à l'observateur avec la meilleure qualité possible. Depuis des années de nombreuses méthodes d'évaluation ont été développées de manière à prédire la qualité perçue par l'observateur. Ces modèles, développés à partir de modélisations plus ou moins fines du système visuel humain prennent en compte les dégradations introduites par le codage du contenu télévisuel et par les différents types de transmission, mais se focalisent rarement sur l'écran. L'écran à tube cathodique a été amélioré des années durant de manière à restituer une image de la meilleure qualité possible. Il est considéré à juste titre comme un écran « transparent », en ce sens qu'il n'apporte pas de dégradations supplémentaires au contenu qu'il affiche, ou tout du moins que ces dégradations sont négligeables vis à vis des dégradations liées au codage ou à la transmission. L'utilisation des écrans à cristaux liquides change quelque peu la donne. On vient de le voir, de nouvelles distorsions spécifiques à ce type d'écran sont introduites. Ces distorsions visuelles ne peuvent plus être négligées car leur impact sur la qualité perçue est important.

Parmi les dégradations citées plus haut, nous avons choisi de nous intéresser au flou de mouvement. Cette dégradation est spécifique aux écrans utilisant un affichage par maintien (matrice active) et peut se révéler très gênante lors de la visualisation de séquences contenant beaucoup de mouvement. Comme nous allons le voir par la suite ce phénomène est avant tout un phénomène perceptif : en effet l'image présentée par l'écran ne présente objectivement aucun flou, le flou apparaît seulement sur la rétine de l'observateur. Il convient donc d'étudier cette distorsion particulière à plusieurs niveaux :

- *Identifier et caractériser le flou de mouvement de manière métrologique.* Il est important de mesurer le flou de mouvement avec des instruments de mesures physiques, de manière à comprendre quelles caractéristiques de l'écran influencent le flou de mouvement, et pour pouvoir ensuite mesurer de manière objective les améliorations apportées par les solutions mises en œuvre.
- *Mesurer et modéliser la perception du flou de mouvement.* En tant que phénomène perceptif, le flou de mouvement doit être étudié en relation avec les propriétés du système visuel humain. Il convient d'étudier avec précision la réponse du SVH à ce type de dégradation de manière

notamment à déterminer des seuils de perception et des seuils de tolérance.

- *Étudier et évaluer l'influence du flou de mouvement sur la qualité globale des images affichées.*

En effet, c'est l'impact final sur la qualité des images et séquences affichées qui nous intéresse réellement. Une fois le flou de mouvement caractérisé de manière photométrique et perceptuelle, il nous faut évaluer dans quelle mesure la qualité visuelle globale est affectée.

La seconde partie de ce mémoire de thèse traite du premier axe d'étude : la caractérisation photométrique du flou de mouvement. La troisième partie est dédiée à l'étude de la perception du flou de mouvement à travers différentes expériences psychophysiques. Enfin, la quatrième et dernière partie concerne l'évaluation subjective de la qualité visuelle sur écrans à cristaux liquides.

2.5 Conclusion

Dans ce chapitre, nous avons présenté le fonctionnement général des écrans à cristaux liquides, pour comprendre la manière dont les images sont affichées sur ce type d'écrans. Il a été montré que la restitution des images sur écrans LCD était très éloignée de la restitution des images sur écran CRT. La principale différence réside dans la méthode d'affichage. Dans un écran CRT, l'adressage des pixels se fait par balayage : un seul pixel est affiché à un instant donné, puis sa luminosité décroît très rapidement. Il reste donc à l'état « éteint » entre deux images. Les écrans LCD utilisent un affichage à adressage matriciel actif : la luminosité des pixels est maintenue durant toute la période de rafraîchissement, jusqu'à ce qu'une nouvelle valeur leur soit adressée. Ils ne passent pas par l'état « éteint » entre deux images.

La technologie d'affichage à cristaux liquides est relativement récente et de nombreuses distorsions visuelles sont encore à déplorer sur les écrans LCD. Parmi ces distorsions la plus importante semble être le flou de mouvement. Cette dégradation est directement liée au maintien de la luminosité durant l'affichage et n'existait pas sur les écrans de type impulsif comme l'écran à tube cathodique. De manière à évaluer son importance sur la qualité globale des images affichées sur les écrans LCD, il convient d'étudier ce phénomène sous différents axes : de manière métrologique tout d'abord, en caractérisant le flou de mouvement à l'aide d'appareils de mesure photométriques (partie 2), de manière perceptuelle ensuite, à travers différentes expériences psychophysiques visant à modéliser la perception visuelle de cette distorsion (partie 3), et enfin de manière plus globale à travers des tests d'évaluation subjective de la qualité visuelle (partie 4).

Deuxième partie

Mesures photométriques du flou de mouvement sur les écrans à cristaux liquides

Chapitre 3

Le flou de mouvement sur les écrans à cristaux liquides

Sommaire

3.1	Introduction	43
3.2	Flou de mouvement	44
3.2.1	Causes	44
3.2.2	Formation du flou de mouvement	45
3.2.3	Méthodes de réduction du flou de mouvement	47
3.3	Réponse temporelle indicielle d'un écran LCD	48
3.3.1	Mesure de la réponse indicielle	49
3.3.2	Le temps de réponse	51
3.3.3	Normalisations concernant le temps de réponse des écrans LCD	54
3.4	Profil spatial d'un contour en mouvement	57
3.4.1	Systèmes de mesures du flou de mouvement par dispositifs imageurs	57
3.4.2	Mesure du flou de mouvement	58
3.4.3	Largeur de flou	61
3.4.4	Valeur globale du flou de mouvement pour un écran	63
3.5	Conclusion	64

3.1 Introduction

Les écrans à cristaux liquides sont encore sujets à de nombreux défauts et des améliorations restent nécessaire pour atteindre une qualité d'affichage du niveau de celle des écrans à tube cathodique. Intrinsèquement adaptés à l'affichage de contenus fixes, les écrans LCD souffrent d'une perte de qualité lorsqu'ils affichent des contenus animés [ITU 05] (typiquement les signaux de télévision sur un téléviseur ou les applications multimédia sur un moniteur informatique). Le responsable de cette perte de qualité est un défaut particulier des écrans LCD que l'on connaît sous le nom de « flou de mouvement » car il présente des similitudes (aussi bien au niveau des causes que des conséquences)

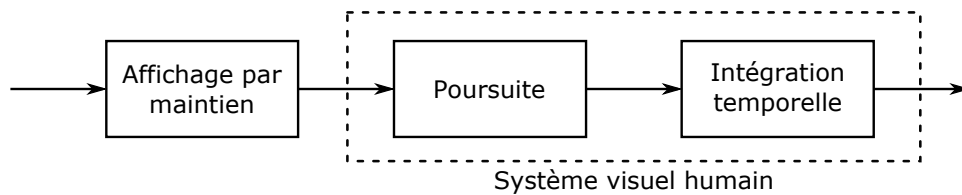


FIGURE 3.1 – Schéma-bloc illustrant les phénomènes mis en jeu dans la formation du flou de mouvement sur les écrans à cristaux-liquides.

avec le flou de mouvement ou flou de bougé présent sur les capteurs. Depuis plusieurs années, de nombreux dispositifs sont mis en œuvre dans les écrans LCD pour réduire cet artefact. Cependant, aucune mesure standard du flou de mouvement n'existe à l'heure actuelle et il peut s'avérer difficile de mesurer la performance de ces dispositifs de réduction. Dans ce contexte, la caractérisation et la mesure du flou de mouvement, à l'aide d'instruments optiques dans un premier temps, constituent un besoin réel pour les fabricants d'écrans.

Dans ce chapitre, nous présentons le flou de mouvement perçu sur les écrans à cristaux liquides. Cette distorsion visuelle est due à la combinaison de l'affichage par maintien propre aux écrans LCD et du mouvement de poursuite réalisé par le SVH lors du suivi d'un objet en mouvement. Nous expliquons ensuite pourquoi la mesure de la réponse indicielle et du temps de réponse de l'écran ne suffit pas à caractériser le flou de mouvement. Enfin, nous passerons en revue les différents protocoles de mesure par dispositifs imageurs et présenterons celui qui a été mis en œuvre durant ces travaux de thèse. Nous introduirons également le formalisme mathématique utilisé, ainsi que les différentes spécifications existantes concernant la mesure du temps de réponse et du flou de mouvement.

3.2 Flou de mouvement

3.2.1 Causes

Le flou de mouvement est un défaut particulier des écrans à cristaux liquides. Il apparaît sur les contours d'un objet en mouvement lorsque l'observateur le suit des yeux. Les différents facteurs responsables de la formation du flou de mouvement sur la rétine sont illustrés par le schéma-bloc de la figure 3.1. Deux causes principales doivent être distinguées [Kurita 01] :

- La première cause est de nature technologique. Comme souligné précédemment, sur un écran à cristaux liquides la luminance d'un pixel est maintenue pendant toute la période d'affichage, on parle alors d'affichage par maintien. Ce comportement est différent de celui des écrans CRT pour lesquels la réponse de chaque pixel consiste en une impulsion lumineuse de durée très brève. On parle alors d'affichage impulsionnel. La figure 3.2 illustre cette différence fondamentale entre les deux types d'écrans.
- La seconde cause est associée au système visuel humain. Lorsque l'on suit des yeux un stimulus en déplacement, un mouvement de type « poursuite » est mis en œuvre par le système visuel humain

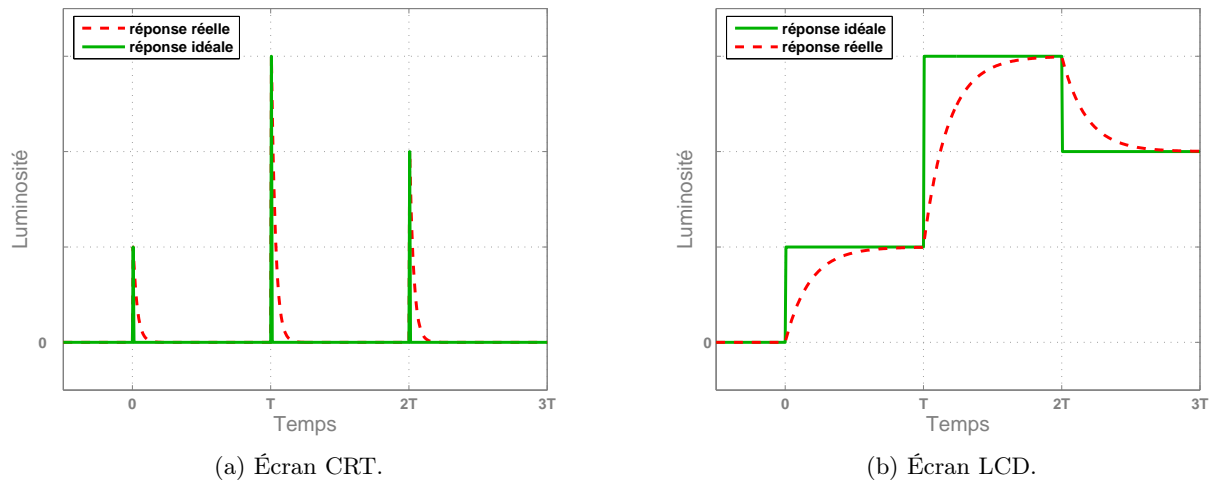


FIGURE 3.2 – Évolution temporelle de la luminance des pixels d'un écran CRT (a) et d'un écran à cristaux liquide (b). La différence entre réponse idéale (trait plein) et réponse réelle (trait pointillé) est due à la décroissance exponentielle de la luminance des luminophore (CRT) et au temps de réactivité non nul des cristaux liquides (LCD).

(cf. A.3.3). Ce type de mouvement permet de stabiliser l'objet en mouvement sur la rétine, de manière à pouvoir l'analyser avec la zone fovéale. La stabilisation n'est pas parfaite car l'œil est sans cesse en train de corriger sa position et sa vitesse. Le contraste du stimulus en mouvement peut également influencer la poursuite. Néanmoins, pour des stimuli très contrastés (de l'ordre de 100 fois le seuil différentiel de visibilité), la poursuite peut être considérée comme réalisée parfaitement [Spering 05].

3.2.2 Formation du flou de mouvement

La figure 3.3 illustre la formation du flou de mouvement sur la rétine. Le stimulus considéré est un contour vertical entre une région de luminance maximale L_{max} à gauche de l'écran et une région de luminance minimale L_{min} à droite de l'écran. Ce contour est animé d'une vitesse horizontale V constante de 10 pixels par image, pour une fréquence d'affichage $f = \frac{1}{T}$. Étant donné le caractère unidimensionnel du stimulus, nous ne considérons qu'une ligne horizontale de 40 pixels sur l'écran. À $t = 0$, le contour est positionné en $x = 0$ et tous les pixels de la ligne situés à une position x telle que $x > 0$ ont une valeur de luminance L_{min} . À $t = T$, le contour est déplacé de 10 pixels vers la droite, il se retrouve positionné en $x = 10$. La luminance des pixels de position x telle que $0 < x \leq 10$ varie donc de L_{min} à L_{max} . Ce changement de luminance n'est pas immédiat à cause de la réactivité des cristaux liquides, l'évolution de la luminance des pixels en fonction du temps est donc similaire à celle décrite figure 3.2b. À $t = 2T$, ce sont les 10 pixels suivants qui sont à leur tour soumis à un changement de luminance de L_{min} à L_{max} et ainsi de suite.

L'observateur suit des yeux le contour en mouvement. Ses yeux sont donc animés d'un mouvement

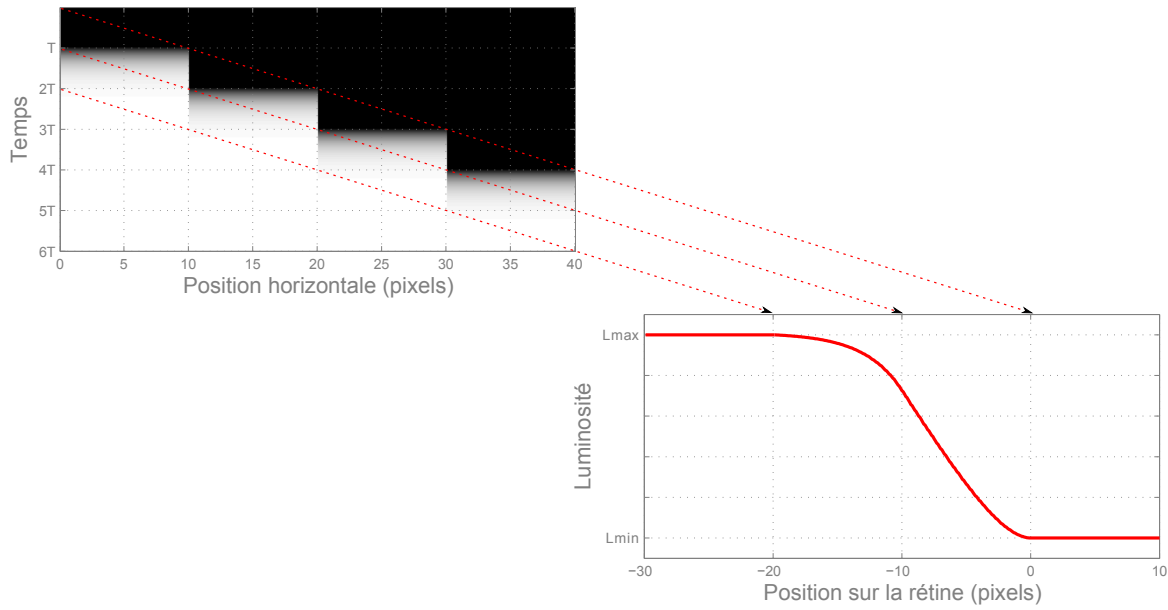


FIGURE 3.3 – Formation du flou de mouvement sur la rétine, dans le cas d'un écran LCD.

de poursuite d'une vitesse théoriquement égale à celle du contour. Si le contour était animé d'un mouvement fluide et continu, il serait donc stabilisé sur la rétine (comme c'est le cas lorsque l'on suit des yeux un objet dans une scène réelle). Cependant, durant une période d'affichage $[kT; (k+1)T]$, le contour est momentanément fixe sur l'écran, tandis que la rétine se déplace. L'image du contour est donc déplacée sur la rétine, ce qui conduit à un flou de bougé. La trajectoire des yeux est représentée sur la figure par les droites pointillées. En première approximation le comportement temporel de l'œil peut être vu comme un intégrateur dont les constantes de temps sont faibles en regard de la période d'affichage [Ahumada 98], l'intégration est donc effectuée de façon complète sur la rétine. Le profil spatial du contour en mouvement peut alors être obtenu en effectuant la moyenne des variations de luminance suivant la direction du mouvement, comme illustré figure 3.3.

À titre de comparaison, la figure 3.4 présente l'évolution spatio-temporelle de la luminance d'une ligne de pixels d'un écran CRT sur lequel est affiché le même stimulus. La réponse temporelle des pixels est similaire à celle décrite figure 3.2a. Contrairement à ce qui se passe sur un écran à cristaux liquides, la luminance des pixels n'est pas maintenue durant la période d'affichage. Par conséquent, lorsque l'œil se déplace pendant l'intervalle de temps entre deux images, il n'y a pas (ou très peu) de lumière intégrée sur la rétine en mouvement. L'image du contour est donc nette sur la rétine.

Cette explication permet d'ores et déjà d'identifier les différents facteurs qui influencent la quantité de flou qui apparaît sur les contours d'un objet en mouvement.

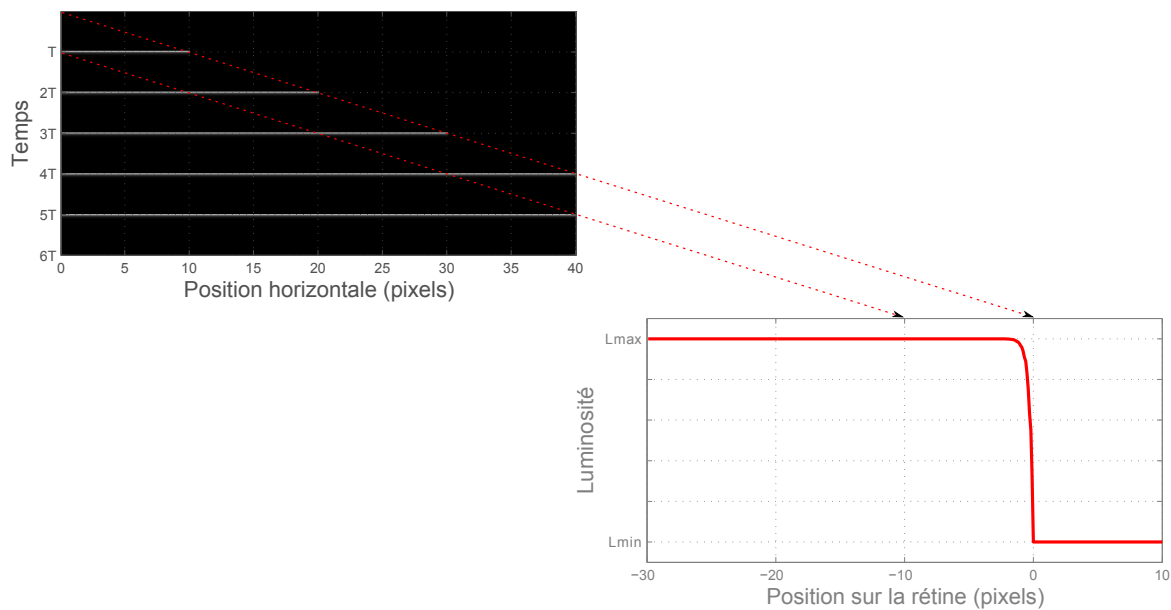


FIGURE 3.4 – Formation du flou de mouvement sur la rétine, dans le cas d'un écran CRT.

- La taille du flou sera d'autant plus importante que la vitesse du mouvement est élevée. Néanmoins ce paramètre est lié au contenu affiché et ne peut pas être réduit.
- La réponse indicielle temporelle des pixels de l'écran LCD joue un rôle essentiel. Cependant le temps de réponse n'est pas le paramètre le plus important : il convient de noter que même dans le cas d'une réponse indicielle idéale, le flou de mouvement serait encore présent. En revanche, la forme de la réponse est à prendre en compte, comme on peut le constater sur les écrans à tube cathodique, une réponse de type impulsionnel ne conduit pas à la formation du flou de mouvement.
- Enfin la durée de la période d'affichage a un impact sur l'extension spatiale du flou. Il est évident au vu de la figure 3.3 que si la période d'affichage est diminuée, la taille du flou sera diminuée d'autant.

3.2.3 Méthodes de réduction du flou de mouvement

Plusieurs méthodes ont été proposées ces dernières années pour réduire le flou de mouvement perçu sur LCD. On peut en distinguer quatre grandes familles :

Le clignotement du rétro-éclairage en synchronisation avec la période d'affichage.

Ce type de dispositif consiste à approcher une réponse de type impulsionnel, comme sur les écrans CRT. Les sources lumineuses permettant l'éclairage de l'écran sont éteintes puis rallumées plusieurs fois par période d'affichage, et en synchronisation avec cette dernière [Fisekovic 01, Feng 06, Shuyterman 06, Song 09]. Cette méthode permet de réduire la quantité de lumière intégrée sur la rétine lorsque celle-ci se déplace sur le contour. Néanmoins, la luminance maximale de l'écran est

également réduite, et le clignotement des sources lumineuses peut être visible si la fréquence de ce clignotement est inférieure à la fréquence de papillotement critique (cf. annexe A.4.2.2), notamment en zone périfovéale.

L'insertion de noir.

Cette méthode est assez similaire à la précédente. Elle consiste à insérer des plages noires dans le flux vidéo, toujours dans l'optique d'approcher une réponse impulsionnelle. Cela peut être fait en insérant une image complètement noire entre chaque image du flux vidéo d'origine (cela nécessite cependant de doubler la fréquence d'affichage de l'écran), ou bien en insérant une bande noire dans les images originales du flux vidéo et en faisant défiler ce bandeau horizontalement d'une image à l'autre [Nose 01, Kim 07]. On observe cependant les mêmes inconvénients concernant la luminance maximale

L'augmentation de la fréquence d'affichage.

Le traitement est réalisé en insérant dans le flux vidéo d'origine des images obtenues par interpolation de deux images originales successives. La période d'affichage est alors divisée par deux, par trois ou par quatre (selon que l'on insère une, deux ou trois images interpolées) et le flou de mouvement est diminué d'autant. Ce processus temps-réel est effectué par un dispositif inclus dans l'écran [Bellers 07, Kim 09]. Il est cependant assez coûteux et une mauvaise interpolation temporelle peut conduire à la formation d'artefacts visibles.

Le filtrage inverse par compensation de mouvement.

Cette technique consiste à traiter l'image pour compenser les effets du flou de mouvement sur celle-ci, en se basant sur la fonction de transfert de modulation de l'écran [Klompener 04, Haiyan 07]. Encore une fois, le traitement est très complexe et doit être effectué en temps réel. Une mauvaise estimation de mouvement peut là aussi conduire à la formation de distorsions visibles.

Malgré quelques modèles disponibles sur le marché (notamment des téléviseurs LCD 200 Hz depuis fin 2008), la plupart de ces traitements sont encore à l'état de développement et d'amélioration. La caractérisation du flou de mouvement d'un écran est donc un outil nécessaire pour pouvoir évaluer les apports de ces dispositifs de réduction.

3.3 Réponse temporelle indicielle d'un écran LCD

La quantité du flou de mouvement sur la rétine dépend fortement de la réponse temporelle indicielle de l'écran à cristaux liquides. Cette réponse indicielle est caractérisée par son temps de montée, ou de descente selon la transition lumineuse considérée. Ce temps est communément appelé temps de réponse, il est lié à la réactivité des cristaux liquides qui composent chaque sous-pixel. Il correspond en fait au temps nécessaire aux cristaux pour changer leur alignement de manière à modifier la transmittance de la cellule et donc la luminance du sous-pixel. Cette section décrit les moyens mis en œuvre pour mesurer cette réponse indicielle.

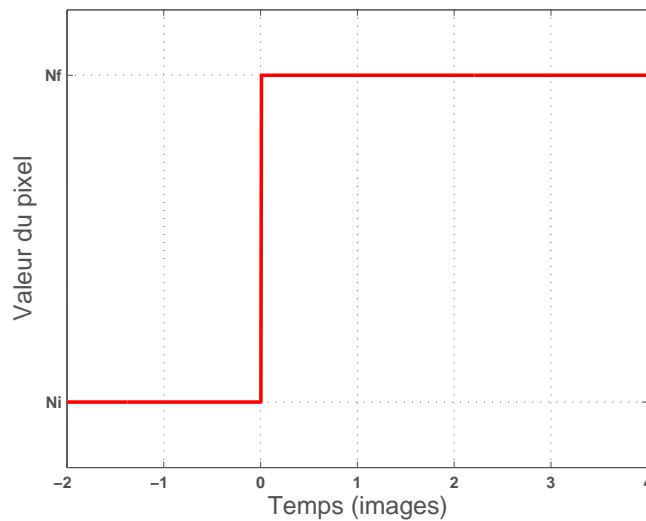


FIGURE 3.5 – Évolution temporelle de la valeur d'un sous-pixel soumis à la transition $N_i \rightarrow N_f$.

La réponse indicielle d'un sous-pixel dépend de la valeur de départ et de la valeur finale de la commande échelon. On fait ici l'hypothèse que tous les sous-pixels d'un écran ont le même comportement temporel. En réalité de faibles différences peuvent apparaître mais celle-ci sont négligeables. Il suffit donc de connaître le comportement d'un sous-pixel pour en déduire le comportement du pixel entier. Pour cette raison il n'est pas utile de mesurer la réponse indicielle entre deux niveaux de couleur puisque cette réponse peut être obtenue en combinant les réponses des trois sous-pixels. Par conséquent, on ne mesurera ici que la réponse indicielle entre deux niveaux de gris.

3.3.1 Mesure de la réponse indicielle

3.3.1.1 Signal de commande

On mesure la réponse indicielle temporelle à un signal échelon entre un niveau de gris initial N_i et un niveau de gris final N_f . Le signal d'entrée est un signal échelon, désigné par le terme *transition* et noté $N_i \rightarrow N_f$, il est défini par les relations suivantes, pour une transition montante et une transition descendante respectivement :

$$N_i \rightarrow N_f(t) = (N_f - N_i) \cdot H(t) + N_i \quad \text{si } N_i < N_f \quad (3.1)$$

$$N_i \rightarrow N_f(t) = (N_f - N_i) \cdot (1 - H(t)) + N_i \quad \text{si } N_i > N_f \quad (3.2)$$

avec $H(t)$ la distribution de Heaviside.

Un exemple de transition montante est donnée figure 3.5.

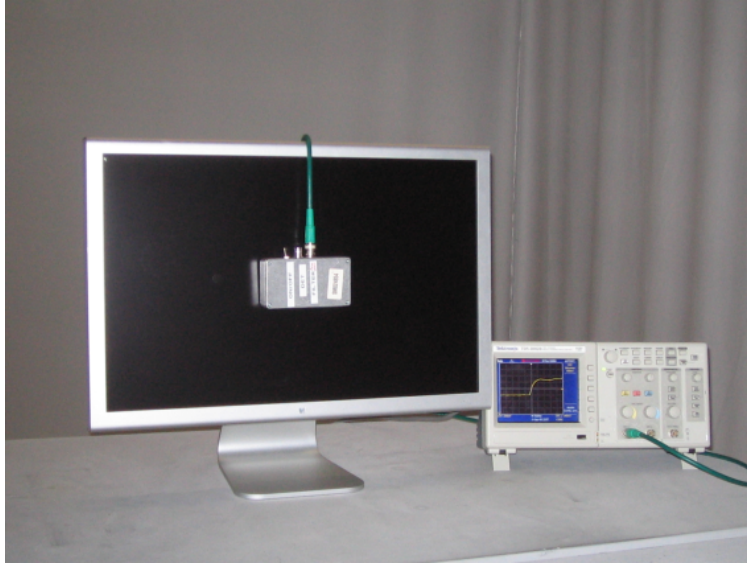


FIGURE 3.6 – Dispositif de mesure de la réponse indicielle d'un écran.

3.3.1.2 Dispositif

Pour mesurer cette réponse indicielle des écrans testés, nous avons utilisé une photo-diode fabriquée par Acreo AB avec qui nous avons collaboré pour réaliser ces travaux. La diode est logée dans un boîtier métallique avec une ouverture de 2 mm par 2 mm, le nombre de pixels correspondant dépend de la taille d'un pixel (celle-ci varie généralement de 0.2 mm pour les moniteurs de bureau jusqu'à 0.6 mm pour les téléviseurs). Un morceau de velours noir est fixé sur le boîtier, à la fois pour protéger l'écran du contact avec la sonde et pour isoler le capteur de toute lumière parasite.

Le dispositif de mesure de la réponse indicielle est présenté figure 3.6. Il est constitué de la photo-diode connectée à un oscilloscope numérique. La fréquence d'échantillonnage du signal en sortie de la photo-diode a été fixée à 10 kHz. Le temps de réponse de la photo-diode est très inférieur à 0.1 ms.

3.3.1.3 Signal mesuré

Les pixels de l'écran sont soumis à une transition $N_i \rightarrow N_f$ entre un niveau de gris initial N_i et un niveau de gris final N_f . La réponse des pixels à cette transition est notée $\mathcal{R}_{N_i \rightarrow N_f}(t)$. Elle correspond à l'évolution temporelle de la luminance du pixel soumis à la transition $N_i \rightarrow N_f$, elle s'exprime en candela par mètre carré (cd/m^2). La luminance émise par un pixel pour un niveau de gris donné N dépend de la fonction électro-optique de l'écran, notée γ (cette fonction donne la valeur de la luminance émise par l'écran pour chaque niveau de gris N_k). Les luminances initiale L_i et finale L_f de la réponse indicielle sont donc :

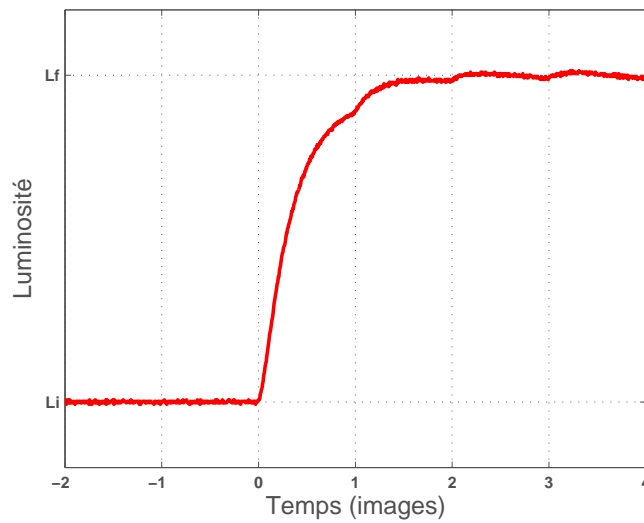


FIGURE 3.7 – Réponse indicielle $\mathcal{R}_{N_i \rightarrow N_f}(t)$ d'un pixel soumis à une transition $N_i \rightarrow N_f$.

$$\begin{aligned} L_i &= \gamma(N_i) \\ L_f &= \gamma(N_f) \end{aligned}$$

La réponse indicielle est ensuite obtenue de la façon suivante :

$$\mathcal{R}_{N_i \rightarrow N_f}(t) = (L_f - L_i) \cdot S(t) + L_i \quad (3.3)$$

où $S(t)$ est le signal en sortie de la photo-diode, normalisé entre 0 et 1.

Un exemple de réponse indicielle est présenté figure 3.7. La forme de la réponse indicielle dépend des caractéristiques intrinsèques des cristaux liquides qui composent chaque sous-pixel. Elle peut être modélisée de différentes manières en étudiant le comportement des cristaux liquides et de l'électronique associée. Sur ce sujet, on peut notamment citer les travaux de Aoki [Aoki 96], McCartney [McCartney 03] et Adam [Adam 07].

3.3.2 Le temps de réponse

Le temps de réponse d'un pixel est défini comme le temps de montée de la réponse indicielle dans le cas d'une transition montante ($N_i < N_f$) et comme le temps de descente de la réponse indicielle dans le cas d'une transition descendante ($N_i > N_f$). Il est fortement dépendant des niveaux de gris N_i et N_f . Il y a donc autant de temps de réponse qu'il y a de transitions $N_i \rightarrow N_f$, soit $2^n(2^n - 1)$ pour un écran avec une profondeur de pixel de n bits (65280 dans le cas d'un écran 8 bits).

3.3.2.1 Théorie

Le temps de réponse du pixel dépend des caractéristiques dynamiques des molécules de cristaux liquides contenues dans le pixel. Lorsque le signal de commande est modifié, les molécules de cristaux liquides changent leur orientation de manière à modifier la transmittance du pixel. Par exemple, dans le cas des cristaux liquides de type nématique¹, on distingue principalement deux types de contraintes : une première contrainte liée au couple de torsion appliqué pour déplacer les molécules de cristal, et une seconde contrainte liée à la résistance à ce mouvement (écoulement dynamique, viscosité) [McCartney 03]. Les forces de résistance au mouvement sont intrinsèques au type de cristal liquide et à son environnement (inclinaison initiale, densité de cellules, température, etc.).

Le couple de torsion appliqué aux cellules de cristal résulte de la somme de deux couples :

- un couple d’excitation, induit par le champ électrique ;
- un couple de retour à l’état initial, induit par la constante de raideur du positionnement des cristaux liquides.

Le temps requis pour déplacer les molécules d’une orientation à l’autre dépend de l’état actuel de la déformation cristalline et diffère selon que le couple induit par le champ électrique est augmenté ou diminué. En effet on distingue deux cas [McCartney 03] :

- La valeur courante du couple d’excitation est supérieure à la valeur précédente : le temps de réponse est alors approximativement inversement proportionnel au carré de la tension appliquée (cela dépend des facteurs – le plus souvent non linéaires – liés aux contraintes de résistance au mouvement) ;
- La valeur courante du couple d’excitation est inférieure à la valeur précédente : le temps de réponse est approximativement inversement proportionnel à la tension appliquée (cela dépend principalement des facteurs passifs, non linéaires, permettant le positionnement des cristaux).

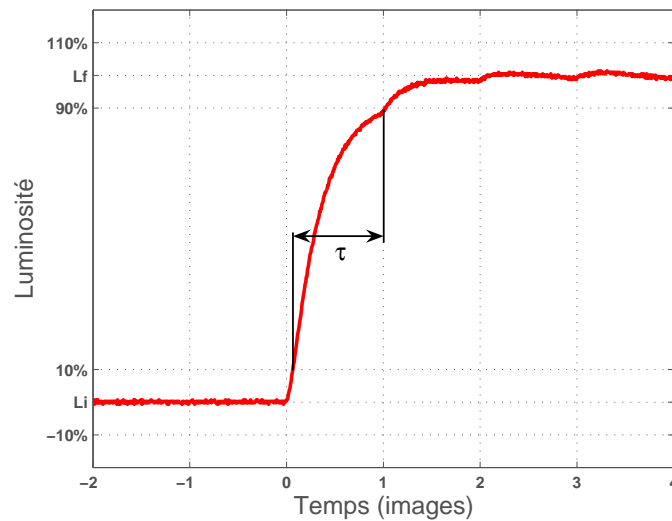
Ces deux comportements nous permettent d’établir une première constatation concernant le temps de réponse :

- Pour les modes de cristaux liquides dont l’état est opaque au repos (VA, IPS), le temps de réponse d’une transition montante sera généralement plus long que le temps de réponse d’une transition descendante.
- Pour les modes de cristaux liquides dont l’état est allumé au repos (TN), le temps de réponse d’une transition descendante sera généralement plus long que le temps de réponse d’une transition montante.

3.3.2.2 Calcul du temps de réponse pour une transition

Le temps de réponse d’un pixel à une transition $N_i \rightarrow N_f$ est mesuré sur la réponse indicielle $\mathcal{R}_{N_i \rightarrow N_f}$. Il est noté $\tau_{N_i \rightarrow N_f}$ et défini comme le temps que met l’intensité lumineuse émise par le pixel pour passer de la luminance initiale L_i (à 10% près) à la luminance finale L_f (à 10% près). En d’autres termes :

1. Les trois modes de cristaux liquides les plus utilisés (TN, IPS et VA) sont tous de type nématique (cf. tableau 2.1).

FIGURE 3.8 – Mesure du temps de réponse τ sur la réponse indicielle.

$$\tau_{N_i \rightarrow N_f} = t_f - t_i \quad (3.4)$$

avec t_i le temps t le plus faible qui vérifie :

$$\mathcal{R}_{N_i \rightarrow N_f}(t) = L_i \pm 0.1 |L_f - L_i| \quad (3.5)$$

et t_f le temps t le plus grand qui vérifie :

$$\mathcal{R}_{N_i \rightarrow N_f}(t) = L_f \pm 0.1 |L_f - L_i| \quad (3.6)$$

La figure 3.8 illustre la mesure du temps de réponse pour une transition montante.

3.3.2.3 Artefacts temporels

Différents types d'artefacts temporels peuvent apparaître sur la réponse indicielle mesurée.

- La réponse peut être bruitée (cf. figure 3.9a). Ce bruit peut être lié à l'acquisition, à la transmittance de la cellule de cristaux liquides, ou encore à la source lumineuse.
- Sur les écrans équipés de tubes fluorescents (cf. section 2.3.2), la luminosité du rétro-éclairage est ajustée par modulation de largeur d'impulsion (MLI). Il en résulte des artefacts périodiques qui peuvent s'avérer gênants si l'on veut mesurer le temps de montée de la réponse indicielle (cf. figure 3.9b).
- Sur certains écrans, des artefacts peuvent être causés par le rafraîchissement de l'image (cf. figure 3.9c). La fréquence de ces artefacts est la même que la fréquence de rafraîchissement de l'image [Becker 07].

- Pour améliorer le temps de réponse des écrans, des techniques de compensation du temps de réponse ont été développées [McCartney 03, Kumar 05]. Elles consistent à appliquer une sur-commande au pixel de manière à accélérer son changement d'état. Cependant, lorsque les paramètres de la sur-commande sont mal ajustés, une sur-intensité lumineuse (ou une sous-intensité dans le cas d'une transition descendante) peut apparaître sur la réponse indicielle (cf. figure 3.9d).

Ces artefacts rendent difficile la mesure du temps de réponse. Différents traitements doivent alors être mis en œuvre pour obtenir une réponse indicielle régulière sur laquelle le temps de réponse est facilement mesurable. Dans le cas du bruit, il suffit de filtrer la réponse indicielle avec un filtre passe-bas avec une haute fréquence de coupure. Différentes techniques existent pour filtrer les modulations du rétro-éclairage [Becker 08a], notamment la moyenne par fenêtre glissante. Cependant, cette technique modifie la pente de la réponse indicielle et augmente le temps de réponse. Il faut donc compenser les effets de ce type de filtrage. Une autre technique consiste à utiliser un filtre réjecteur de bande (*notch filter*), en prenant soin également de compenser les effets potentiels sur la réponse. Concernant la sur-commande, elle est généralement considérée comme une caractéristique à part entière de la réponse indicielle. Elle est donc prise en compte dans la mesure du temps de réponse (par exemple, si la surintensité est telle que le profil de luminance dépasse 110% de la dynamique de la transition). On mesure également l'intensité de la surintensité [VESA 05].

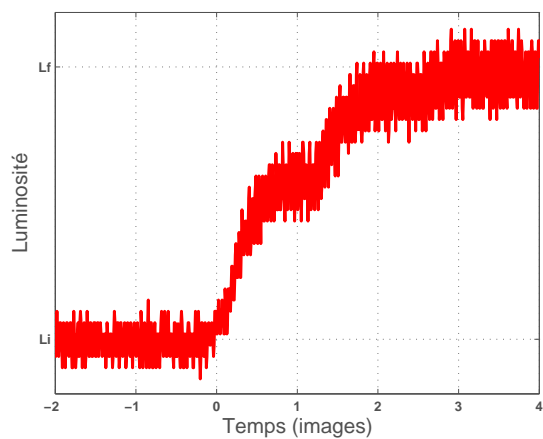
3.3.3 Normalisations concernant le temps de réponse des écrans LCD

3.3.3.1 Norme ISO

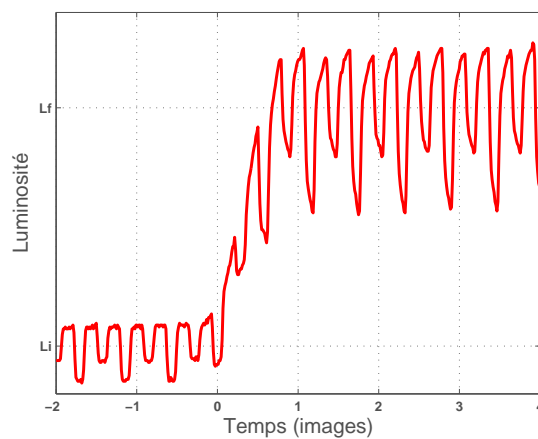
Il existe une norme ISO concernant la mesure du temps de réponse d'un écran LCD : la norme ISO 13406-2 [ISO 01] intitulée *Exigences ergonomiques pour travail sur écrans de visualisation à panneau plat*. Elle spécifie que le temps de réponse d'un écran plat doit être calculé en additionnant le temps de réponse à la transition montante de noir vers blanc $0 \rightarrow (2^n - 1)$ et le temps de réponse à la transition descendante de blanc vers noir $(2^n - 1) \rightarrow 0$ (avec n la profondeur de pixels en bits). En d'autres termes, le temps de réponse ISO τ_{ISO} d'un écran 8 bits est défini par :

$$\tau_{ISO} = \tau_{0 \rightarrow 255} + \tau_{255 \rightarrow 0} \quad (3.7)$$

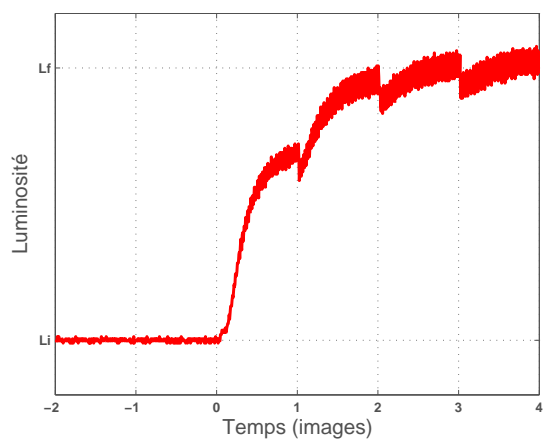
Ce type de mesure a été défini en premier lieu pour les moniteurs informatiques et pour une utilisation de type bureautique. Elle ne s'applique plus dans le cas des téléviseurs ou des moniteurs de bureau à usage multimédia. En effet, la transition de noir-à-blanc, qui équivaut au passage d'une transmittance de 0 à 100% des cristaux liquides, est la transition la plus rapide mais certainement pas le mode de réponse le plus fréquent lors de l'affichage d'un contenu naturel. Les transitions gris-à-gris, plus fréquentes, sont cependant beaucoup plus lentes. L'évaluation de la réactivité d'un écran ne peut donc seulement être faite en tenant compte du temps de réponse ISO.



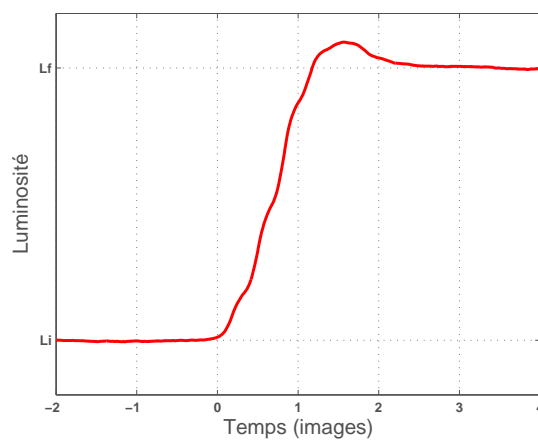
(a) Bruit.



(b) Modulation de largeur d'impulsion.



(c) Rafraîchissement d'image.



(d) Overdrive.

FIGURE 3.9 – Exemples d'artefacts temporels qui peuvent apparaître sur la réponse indicielle d'un pixel.

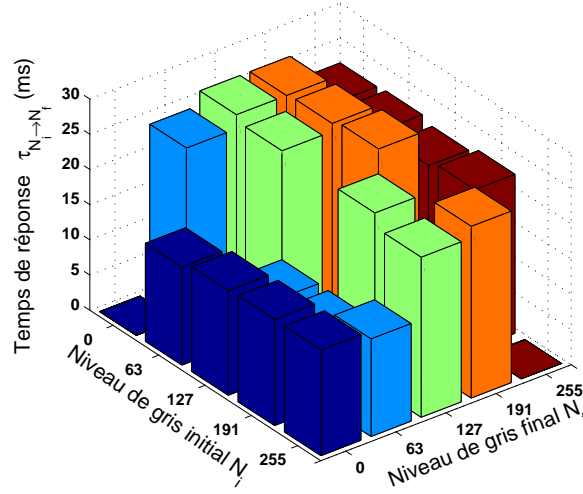


FIGURE 3.10 – Cartographie du temps de réponse d'un ensemble de transitions.

3.3.3.2 Temps de réponse gris-à-gris (VESA)

De manière à obtenir une mesure plus représentative du temps de réponse d'un écran, une alternative a été proposée par la VESA (*Video Electronics Standards Association*) [VESA 05], également spécifiée par une norme ISO plus récente que la précédente [ISO 08]. Cette méthode prend en compte le temps de réponse d'un ensemble de transitions gris-à-gris et non plus seulement celui de deux transitions extrêmes.

Soit \mathbb{T} l'ensemble des transitions $N_i \rightarrow N_f$ choisies, le temps de réponse gris-à-gris (également appelé temps de réponse *gray-to-gray* ou *GTG*), noté τ_{GTG} , est obtenu en effectuant la moyenne des temps de réponse de chaque transition :

$$\tau_{GTG} = \frac{1}{\text{Card}(\mathbb{T})} \sum_{N_i \rightarrow N_f \in \mathbb{T}} \tau_{N_i \rightarrow N_f} \quad (3.8)$$

Bien entendu, cette méthode dépend fortement du choix de l'ensemble \mathbb{T} ainsi que de sa taille. Il est recommandé d'utiliser au moins 5 niveaux de gris uniformément distribués selon le modèle de luminance perçue de l'espace de couleur CIE $L^*a^*b^*$ (cf. section A.3.1.3). Bien que cette mesure du temps de réponse permette d'obtenir une seule valeur, il est assez courant de rencontrer une représentation graphique en trois dimensions permettant de lire la valeur du temps de réponse pour chaque transition, comme illustré figure 3.10 pour un écran IPS. On retrouve d'ailleurs le constat fait dans la section 3.3.2.1 : le temps de réponse des transitions descendantes (au premier plan) sont globalement inférieures à ceux des transitions montantes (au second plan) pour ce type d'écran.

3.3.3.3 Validité du temps de réponse

Le temps de réponse a longtemps été considéré comme un indice fiable de mesure de la qualité de reproduction du mouvement d'un écran. Il reste à l'heure actuelle, le seul indice disponible sur les spécifications techniques des écrans commercialisés. Pourtant, comme on l'a vu précédemment (cf. section 3.2.2), même avec une réponse idéale et un temps de réponse nulle, le flou de mouvement serait toujours présent. En fait, étant donnée la valeur moyenne du temps de réponse des écrans actuels, la contribution du temps de réponse au flou de mouvement est de l'ordre de 20%, les 80% restant étant dus à l'affichage par maintien à proprement parler [Pan 05].

Néanmoins, un temps de réponse assez faible est nécessaire pour s'affranchir des problèmes de rémanence qui apparaissaient sur les premières générations d'écrans à cristaux liquides. Ce problème était caractérisé par le fait qu'une image restait légèrement visible (image fantôme) pendant les quelques images suivantes. Les nouvelles générations d'écrans se sont affranchies de ce défaut grâce à l'utilisation de système de compensation du temps de réponse. Un temps de réponse plus faible que la période d'affichage est donc souhaitable de ce point de vue là. Et cette caractéristique est contrôlée pour obtenir certaines certifications. Par exemple, pour obtenir le label TCO Media Displays [TCO 06] proposé par l'organisme de certification TCO Development, il faut que les temps de réponse gris-à-gris mesurés sur l'ensemble des transitions $\mathbb{T} = N_i \rightarrow N_f$ avec $(N_i, N_f) \in \{0, 63, 127, 191, 255\}^2$ et $N_i \neq N_f$ soient tous inférieurs à 13 ms. Mais il est bien spécifié dans le document présentant cette certification que « le temps de réponse n'est pas une condition suffisante pour assurer une bonne reproduction du mouvement ».

3.4 Profil spatial d'un contour en mouvement

La mesure du temps de réponse ne permet pas de caractériser le flou de mouvement perçu sur un écran LCD. En effet, si une réponse indicielle plus rapide permet de réduire légèrement le flou de mouvement perçu, c'est bien la forme de la réponse qui conditionne la formation du flou sur la rétine de l'observateur. Dans ces conditions, il est nécessaire de trouver une méthode pour mesurer directement le flou de mouvement. Cela peut être réalisé par l'acquisition d'un contour en mouvement à travers un dispositif imageur.

3.4.1 Systèmes de mesures du flou de mouvement par dispositifs imageurs

La mesure du flou de mouvement est effectuée en imitant le comportement de poursuite du système visuel humain. Cette poursuite est réalisé par une caméra électronique qui capture l'image d'un stimulus en mouvement. On distingue deux types de dispositifs :

- Les dispositifs imageurs qui suivent le mouvement du stimulus. Ils sont constitués d'une caméra électronique, le suivi de mouvement est réalisé en montant la caméra sur un rail motorisé ou sur un bras mécanique, ou bien en utilisant un miroir en rotation, la caméra est alors fixe [Oka 04, Someya 07].
- Les dispositifs imageurs constitués d'une caméra stationnaire haute-vitesse. Les caméras utilisées ont une fréquence d'acquisition très élevée et capture plusieurs images pour chaque période de

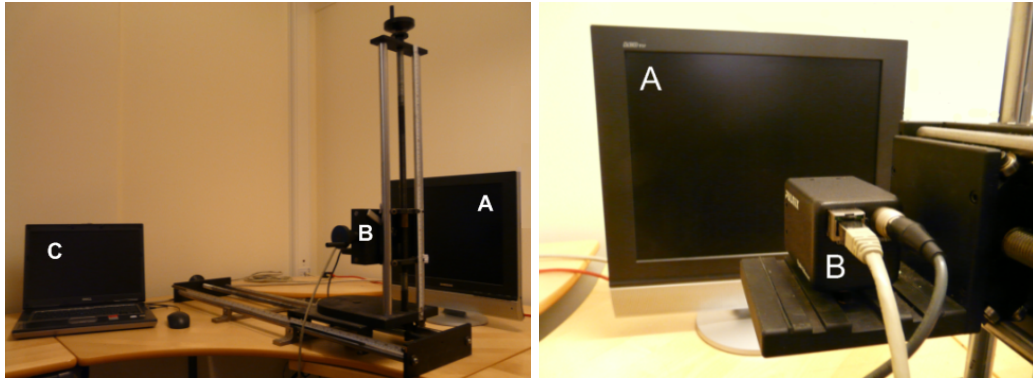


FIGURE 3.11 – Dispositif expérimental : écran testé (A), caméra rapide (B), et ordinateur (C) contrôlant la caméra et générant le stimulus de test.

rafraîchissement de l'écran. Le suivi du mouvement est ensuite simulé en effectuant une compensation du mouvement sur les images capturées [Yamamoto 06, Feng 06].

Pour des raisons pratiques, nous avons préféré mettre en œuvre la seconde méthode pour caractériser le flou de mouvement sur nos écrans. En effet, cette méthode nécessite un dispositif moins complexe en terme de montage, et également moins coûteux. À coût égal, cela permet également d'utiliser une caméra de meilleure qualité permettant d'atteindre une fréquence d'acquisition élevée, améliorant ainsi la résolution temporelle de l'ensemble.

3.4.2 Mesure du flou de mouvement

Dans cette section, nous présentons une méthode de capture du flou de mouvement basée sur l'utilisation d'une caméra rapide. Ce type de méthode est largement utilisé dans l'état de l'art et ne constitue pas une innovation. Néanmoins, les résultats obtenus par ce biais nous serviront de référence par la suite.

3.4.2.1 Dispositif

Le dispositif expérimental utilisé pour effectuer la capture d'un contour en mouvement est représenté figure 3.11. Il est constitué de l'écran test (A), d'une caméra rapide (B) et d'un ordinateur de contrôle qui permet à la fois le contrôle de la caméra et la génération du stimulus affiché sur l'écran test. La caméra rapide utilisée est une caméra CCD JAI PULNiX, modèle TM-6740GE (cf. figure 3.12). Elle est connectée à l'ordinateur de contrôle via une interface Ethernet utilisant le standard GigE Vision. Ce standard, basé sur la technologie Gigabit Ethernet, permet d'atteindre un transfert d'images très rapide.



FIGURE 3.12 – Caméra rapide JAI PULNiX TM-6740GE.

3.4.2.2 Acquisition

La caméra est équipée d'un capteur monochrome CCD Kodak de 640×480 pixels. L'acquisition des données se fait par un balayage progressif, à cette résolution le débit d'images maximum est de 200 images par seconde. Pour étudier l'évolution de la luminance des pixels au cours du temps pendant la durée d'affichage d'une image sur l'écran, il faut utiliser un débit d'acquisition tel qu'à chaque image affichée sur l'écran correspondent plusieurs images capturées par la caméra. La fréquence de rafraîchissement usuelle d'un écran LCD est de 60 Hz pour un moniteur de bureau et de 50/60 Hz (selon les régions du monde) pour un téléviseur. L'augmentation de la cadence d'acquisition est réalisée en utilisant un balayage partiel de la matrice des cellules photo-sensibles. Ce mode d'acquisition permet de réduire la zone utile du capteur à une zone centrale de 224×160 pixels. Le balayage de cette zone réduite étant plus rapide, la cadence d'acquisition peut être augmentée jusqu'à une valeur théorique de 1250 image par seconde. De manière à ce que le nombre d'images acquises pendant la durée d'affichage d'une image à l'écran soit entier, une cadence d'acquisition de 1200 images par seconde a été choisie. Pour un écran avec une fréquence de rafraîchissement de 60 Hz (ce qui sera le cas pour tous les écrans testés ici, sauf mention contraire), 20 images sont acquises pour chaque image affichée à l'écran.

Un dispositif de réglage permet de positionner et d'orienter la caméra de façon très précise. Nous avons choisi une distance d'acquisition telle que chaque pixel de l'écran testé soit capturé par un bloc de 4×4 pixels sur la matrice CCD de la caméra. La zone de l'écran capturée par la caméra a donc une taille de 56×40 pixels, au centre du panneau.

3.4.2.3 Stimuli

Les stimuli ont été générés en utilisant l'extension PsychToolbox de Matlab [Brainard 97]. Celle-ci permet de contrôler directement le contrôleur graphique à l'aide d'une bibliothèque de fonctions adaptées. Pour cette série de mesures, le stimulus consiste en un contour entre deux zones uniformes de niveaux de gris N_i (à droite du contour) et N_f (à gauche du contour). Le contour se déplace de la gauche vers la droite avec une vitesse fixe V .

3.4.2.4 Profil spatial du contour

La figure 3.13 schématise le système de mesure dans son ensemble. Le stimulus se déplace horizontalement sur l'écran avec une vitesse V en pixels par image. En d'autres termes, le stimulus est stationnaire pendant la période de rafraîchissement de l'écran, puis il est décalé de V pixels sur l'image suivante et ainsi de suite.

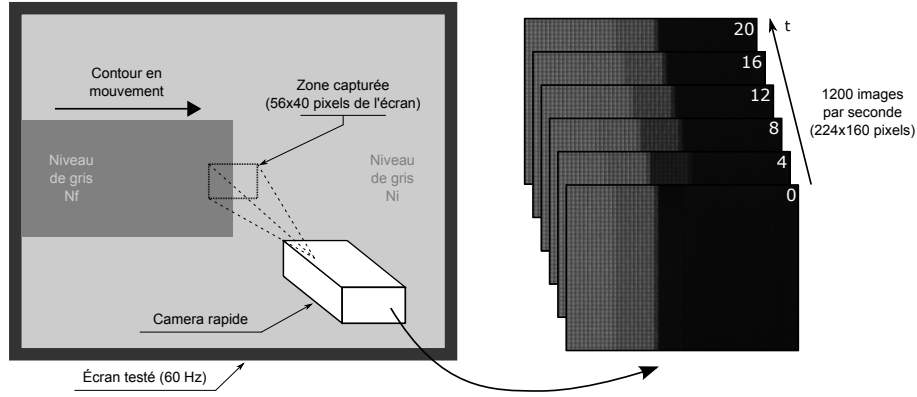


FIGURE 3.13 – Schéma récapitulatif du dispositif de capture d'un contour en mouvement.

En conséquence, les vingt images I_k prises par la caméra pendant une période de rafraîchissement donnée contiennent le contour à la même position. D'un point de vue temporel cependant, l'intensité des pixels varie suivant la réponse indicielle de la transition concernée. La figure 3.14 donne un exemple de 5 images acquises pendant la période de rafraîchissement de l'écran.

De manière à générer le flou de mouvement sur le contour, le déplacement du capteur est simulé. Nous nous plaçons dans l'hypothèse d'un suivi en mouvement idéal, comme décrit dans la section 4.2.2. Chaque image $I_k(x, y)$ ($k = 1 \dots 20$) acquise pendant la période de rafraîchissement de l'écran est décalée spatialement avec un décalage d_k (en pixels) tel que :

$$d_k = \frac{k-1}{20} \cdot V \quad (3.9)$$

avec V la vitesse de déplacement du contour.

Les images décalées sont ensuite additionnées pour simuler l'intégration temporelle du capteur. Le résultat correspond à l'image du contour en mouvement telle qu'elle apparaît sur le capteur en mouvement :

$$I_R(x, y) = \sum_k I_k(x - d_k, y) \quad (3.10)$$

Le profil spatial du contour $\mathcal{C}(x)$ correspond alors à une section de l'image obtenue, le long de la direction du mouvement, prise pour n'importe quelle position verticale y_0 :

$$\mathcal{C}_{N_i \rightarrow N_f}(x) = I_R(x, y_0) \quad (3.11)$$

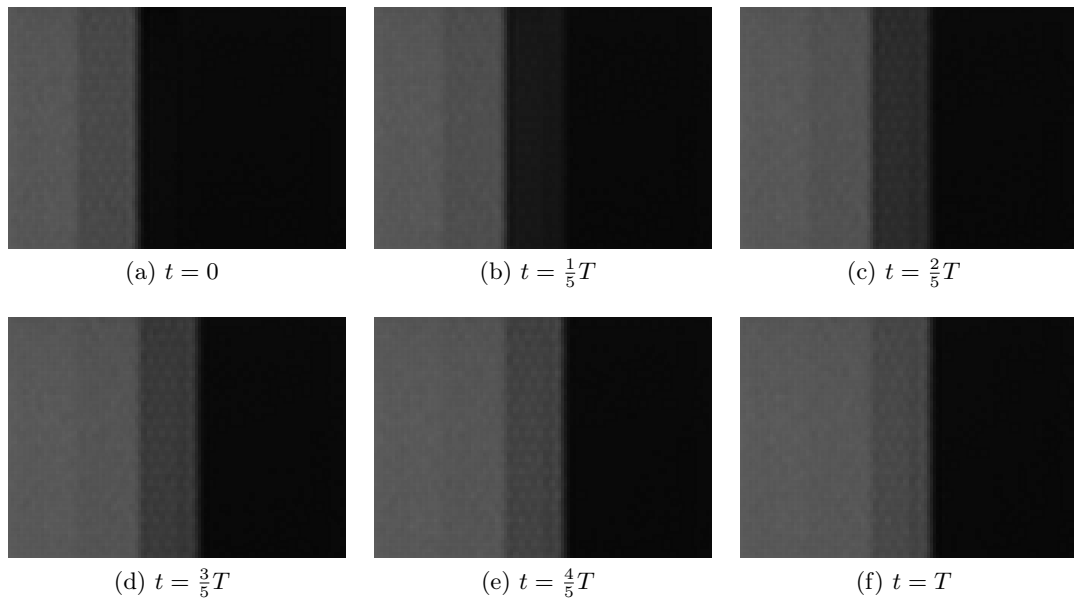


FIGURE 3.14 – Exemples d'images acquises par la caméra durant la période de rafraîchissement de l'écran ($t \in [0; T]$).

Un exemple d'image du contour obtenue par cette méthode est présenté figure 3.15, ainsi que le profil du contour extrait de cette image.

3.4.3 Largeur de flou

La valeur du flou pour une transition donnée $N_i \rightarrow N_f$ est ensuite obtenue en mesurant la largeur du profil spatial du contour flou $\mathcal{C}_{N_i \rightarrow N_f}(x)$. La largeur de flou est notée $BEW_{N_i \rightarrow N_f}$ (*Blurred Edge Width*) et définie comme l'extension spatiale entre le niveau de luminance L_i (à 10% près) à le niveau de luminance L_f (à 10% près). En d'autres termes :

$$BEW_{N_i \rightarrow N_f} = x_f - x_i \quad (3.12)$$

avec x_i la position x la plus faible qui vérifie :

$$\mathcal{C}_{N_i \rightarrow N_f}(x) = L_i \pm 0.1 |L_f - L_i| \quad (3.13)$$

et x_f la position x la plus grande qui vérifie :

$$\mathcal{C}_{N_i \rightarrow N_f}(x) = L_f \pm 0.1 |L_f - L_i| \quad (3.14)$$

La valeur de flou obtenue dépend de la vitesse de déplacement du contour. On utilise par conséquent un autre indicateur, le temps de flou BET (*Blurred Edge Time*). Il est obtenu en normalisant la largeur de flou par la vitesse de déplacement V du contour en mouvement :

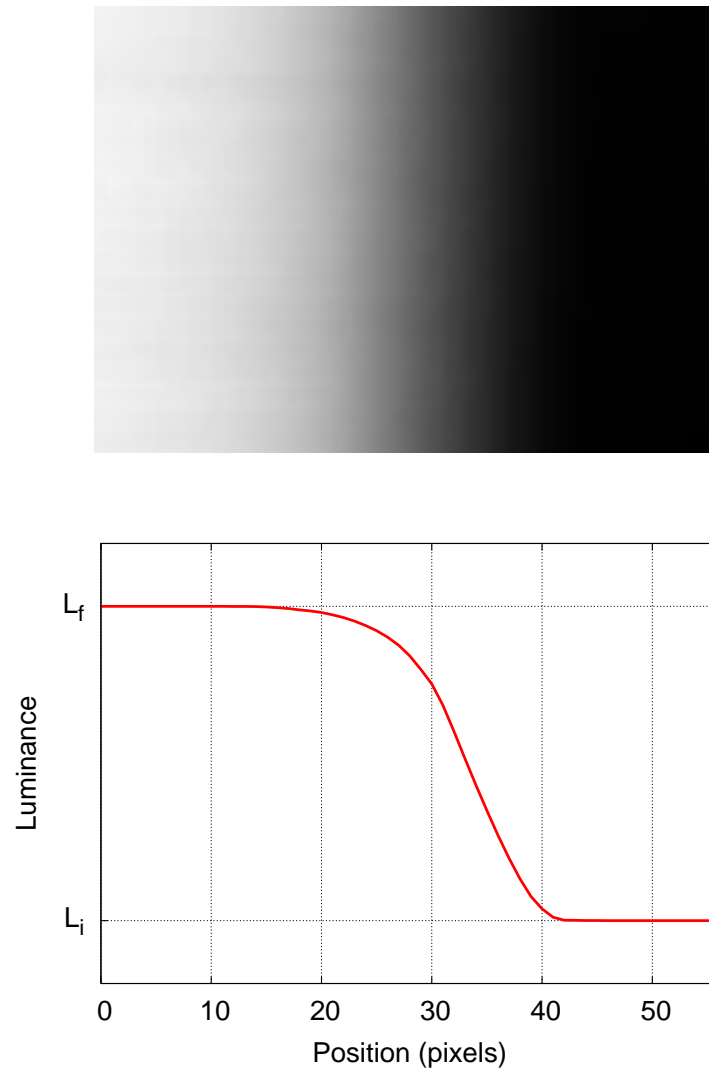


FIGURE 3.15 – Image du contour obtenue après recalage des images acquises par la caméra et intégration temporelle, et profil du contour extrait de cette image.

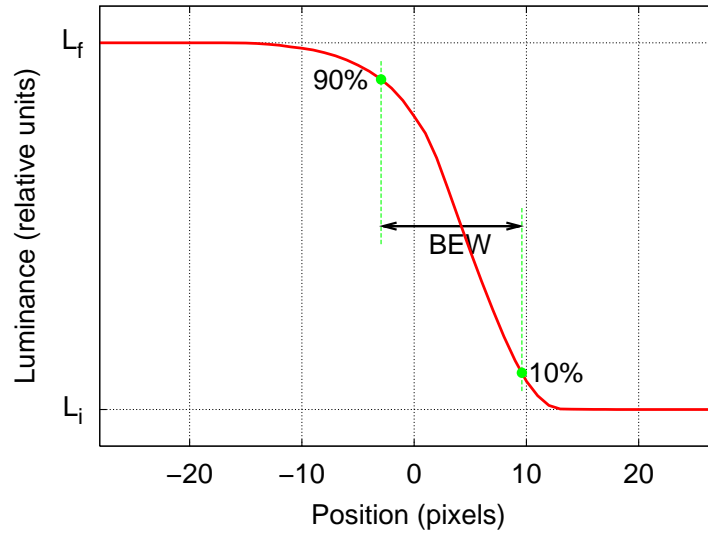


FIGURE 3.16 – Mesure de la largeur de flou BEW sur le profil spatial du contour en mouvement.

$$BET_{N_i \rightarrow N_f} = \frac{BEW_{N_i \rightarrow N_f}}{V} \quad (3.15)$$

3.4.4 Valeur globale du flou de mouvement pour un écran

Il n'existe à l'heure actuelle qu'une seule spécification concernant le calcul d'une valeur globale de la mesure du flou de mouvement perçu sur les écrans à cristaux liquides. Elle se trouve dans le document *Flat Panel Measurement Display Standard* de l'association VESA [VESA 05]. Le groupe de travail FPDM qui a rédigé ce document recommande de mesurer le flou de mouvement d'une manière semblable à celle décrite précédemment pour le temps de réponse (cf. section 3.3.3.2). Il suggère donc de calculer la moyenne des temps de flou BET sur un ensemble T de N transitions. La valeur obtenue est notée MB_{VESA} :

$$MB_{VESA} = \frac{1}{N \cdot V} \sum_{N_i \rightarrow N_f \in T} BEW_{N_i \rightarrow N_f} \quad (3.16)$$

$$MB_{VESA} = \frac{1}{N} \sum_{N_i \rightarrow N_f \in T} BET_{N_i \rightarrow N_f} \quad (3.17)$$

Cette spécification est cependant sujette à controverse. En effet, contrairement au temps de réponse, le flou de mouvement est avant tout un phénomène perceptif. Par conséquent, effectuer la moyenne du flou de mouvement mesuré sur différentes transitions n'a de sens que si la perception du flou de

mouvement est constante sur toutes ces transitions, or rien n'indique que ce soit le cas. Le système visuel humain est sensible au contraste et la perception du contour varie vraisemblablement en fonction des valeurs de luminance L_i et L_f .

Il existe un besoin réel de quantifier la perception du flou de mouvement en regard de la transition impliquée dans le contour, de la vitesse du contour, et du type de réponse indicielle de l'écran, de manière à pouvoir pondérer la moyenne de ces mesures pour obtenir une valeur globale de flou de mouvement qui soit perceptuellement valide.

3.5 Conclusion

Ce chapitre a présenté les éléments clefs permettant de comprendre la formation du flou de mouvement sur les écrans LCD. Cette distorsion visuelle est due à la combinaison d'une première caractéristique typique de la technologie LCD (affichage par maintien) et d'une seconde caractéristique liée au comportement du SVH avec les objets en mouvement (mouvement de poursuite). Les moyens de réduire le flou de mouvement sur LCD sont connus et identifiés. Néanmoins, il est nécessaire de pouvoir caractériser le flou de mouvement de manière fiable pour mesurer l'intérêt d'un dispositif de réduction en regard de son coût.

La mesure de la réponse indicielle permet d'évaluer l'affichage par maintien propre à chaque écran à cristaux liquides. Le temps de réponse des pixels de l'écran, pour une transition donnée, est calculé sur cette réponse indicielle. Néanmoins, cette valeur n'est pas un indicateur suffisant pour caractériser le flou de mouvement, car même avec une réponse idéale (temps de réponse nul), 80% du flou de mouvement serait toujours présent. Pour mesurer le flou de mouvement proprement dit, plusieurs méthodes, basées sur l'utilisation de dispositif imageurs, existent. Nous avons choisi d'utiliser une méthode décrite dans la littérature et qui consiste à simuler le comportement de poursuite du SVH à l'aide d'une caméra rapide. Le profil spatial d'un contour en mouvement peut ainsi être obtenu et le flou de mouvement mesuré. Ce chapitre a également permis de définir les différentes mesures existantes et les termes que nous utiliserons par la suite.

La méthode de mesure utilisée est cependant assez complexe à mettre en place. Dans l'optique de définir un protocole de mesure simple à mettre en œuvre et facilement reproductible d'un laboratoire à l'autre dans des conditions de mesure similaires, nous avons décidé de définir une autre méthode de mesure du flou de mouvement. Cette méthode analytique, qui trouve son origine dans l'analyse de la formation du flou de mouvement, est décrite dans le chapitre suivant.

Chapitre 4

Contribution à la mesure du flou de mouvement sur LCD

Sommaire

4.1	Introduction	65
4.2	Analyse de la formation du flou de mouvement sur le capteur	66
4.2.1	Signal affiché	67
4.2.2	Déplacement du capteur	67
4.2.3	Intégration temporelle	67
4.2.4	Conséquence	69
4.3	Mesures du flou de mouvement	69
4.3.1	Écrans testés	69
4.3.2	Mesures	71
4.4	Résultats	73
4.4.1	Observations	75
4.4.2	Comparaison des résultats obtenus par chacune des deux méthodes	76
4.4.3	Relation entre temps de réponse et temps de flou	78
4.5	Conclusion	79

4.1 Introduction

Dans le chapitre précédent, les méthodes usuelles de mesure du flou de mouvement ont été décrites. Ces protocoles de mesure, basés sur l'acquisition d'un contour en mouvement par un dispositif imageur, ont cependant des défauts majeurs :

- le dispositif est difficile à mettre en place et doit être positionné avec une grande précision ;
- ils sont difficilement reproductibles d'un laboratoire à l'autre, dans les mêmes conditions ;
- ils nécessitent un matériel assez coûteux.

Cette partie des travaux présentés dans ce mémoire a été réalisée au sein du laboratoire de recherche Acreo AB, dans le cadre d'un projet piloté par l'organisme de certification TCO Development.

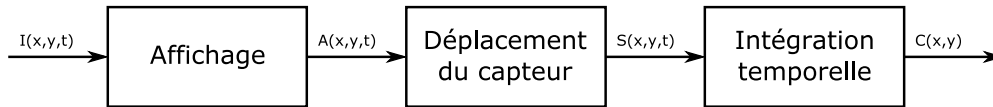


FIGURE 4.1 – Schéma-bloc de la formation du flou de mouvement. Le signal d'entrée $I(x, y, t)$ est affiché par l'écran, il en résulte le signal vidéo $A(x, y, t)$. Celui-ci est alors suivi par le capteur en mouvement pour donner un signal vidéo stabilisé $S(x, y, t)$. Le capteur effectue alors une intégration temporelle qui conduit à une image stationnaire $C(x, y)$.

L'objectif de ce projet fut de définir un protocole de mesure du flou de mouvement qui soit aisé à mettre en œuvre et, surtout, facilement reproductible d'un laboratoire à un autre. Le but à terme étant de déployer ce protocole à l'ensemble des laboratoires de tests de TCO de manière à intégrer une mesure de flou de mouvement dans la prochaine version du label TCO Media Displays [TCO 06].

Dans ce chapitre, les résultats de ce projet de recherche sont présentés. La formation du flou de mouvement sur le capteur (la rétine le cas échéant) est analysée de manière à simplifier le processus de mesure. Cette analyse nous permet de définir une méthode de mesure du flou de mouvement basée sur la réponse indicielle de l'écran testé. Ce protocole est mis en œuvre sur quatre écrans, et les résultats des mesures sont comparés avec ceux obtenus par la méthode classique utilisant la caméra rapide. Enfin, les avantages et défauts de chaque méthode seront présentés.

4.2 Analyse de la formation du flou de mouvement sur le capteur

La formation du flou de mouvement sur un capteur en mouvement (œil ou caméra) est le résultat de la combinaison de deux causes :

- L'intensité lumineuse émise par un pixel est maintenue durant toute la période d'affichage. À chaque changement d'image, cette intensité varie selon une réponse temporelle plus ou moins rapide qui dépend de la transition concernée. La forme de cette réponse ainsi que sa rapidité sont des facteurs influençant le flou de mouvement.
- Le déplacement du capteur est continu alors que les objets en mouvement, dans la scène affichée à l'écran, sont immobile durant la période d'affichage. Un flou de bougé est donc créé sur le capteur en déplacement.

La figure 4.1 illustre les différentes étapes de la formation du flou sur le capteur. Le signal d'entrée $I(x, y, t)$ est reçu par l'écran qui affiche un signal vidéo $A(x, y, t)$ dont les caractéristiques spatio-temporelles dépendent de la fonction d'affichage de l'écran. Le signal arrive ensuite sur le capteur que nous supposons se déplacer à la vitesse de l'objet en mouvement. Le signal vidéo $S(x, y, t)$ à l'entrée du capteur est donc stabilisé. Le signal en sortie du capteur dépend ensuite de la réponse spatio-temporelle de celui-ci. Cette réponse peut être simplifiée en une intégration temporelle.

Nous allons décrire plus précisément chacune de ces étapes.

4.2.1 Signal affiché

Le signal d'entrée consiste en un contour animé d'un mouvement de translation horizontale tel que celui décrit dans le chapitre 3. Le signal affiché est représenté sur la figure 3.3. Ici encore on ne s'intéressera qu'à la dimension horizontale du signal étant donné qu'il n'y a aucune variation selon la direction verticale.

La luminance des pixels situés sur un intervalle spatial $[kV + 1, \dots, (k + 1)V]$ suit une évolution temporelle correspondant à la réponse indicielle à la transition considérée $\mathcal{R}_{N_i \rightarrow N_f}(t)$ (cf. par exemple la réponse indicielle illustrée figure 3.7) avec un décalage temporel égal au nombre d'images k nécessaires pour que le contour atteigne cette zone. Le signal affiché $A(x, t)$ peut donc s'écrire :

$$A(x, t) = \mathcal{R}_{N_i \rightarrow N_f}(t - kT), \quad \forall x \in [kV + 1, \dots, (k + 1)V] \quad (4.1)$$

Ce qui peut être réécrit pour toute position x :

$$A(x, t) = \mathcal{R}_{N_i \rightarrow N_f}\left(t - \left\lfloor \frac{x}{V} \right\rfloor T\right), \quad \forall x$$

avec T la période d'affichage de l'écran et $\lfloor x \rfloor$ la fonction qui renvoie le plus grand entier inférieur ou égal à x . La figure 3.3 donne un exemple de représentation spatio-temporelle du signal affiché par l'écran, pour une vitesse $V = 10$ pixels par image.

4.2.2 Déplacement du capteur

Le flou de mouvement n'apparaît que lorsque le capteur se déplace. C'est notamment le cas lorsqu'un observateur suit des yeux un objet en mouvement. En effet, le système visuel humain possède la capacité de suivre parfaitement un objet se déplaçant à une vitesse constante. Ce suivi stabilise le contour en mouvement sur le capteur. Le contour stabilisé en entrée du capteur $S(x, t)$ est ici reproduit par une compensation de mouvement sur le signal affiché par l'écran $A(x, t)$:

$$S(x, t) = A\left(x + \frac{t}{T}v, t\right) \quad (4.2)$$

$$S(x, t) = \mathcal{R}_{N_i \rightarrow N_f}\left(t - \left\lfloor \frac{x}{V} + \frac{t}{T} \right\rfloor T\right) \quad (4.3)$$

Pour toute position x , ce signal est périodique dans le temps avec une période égale à la période de rafraîchissement de l'écran T . Ce signal est représenté dans un domaine spatio-temporel sur la figure 4.2, pour un contour se déplaçant avec une vitesse de 10 pixels par image.

4.2.3 Intégration temporelle

La dernière étape permet d'obtenir le profil spatial du contour. Pour ce faire on considère en première approximation que le capteur effectue une intégration temporelle complète du signal incident $S(x, t)$.

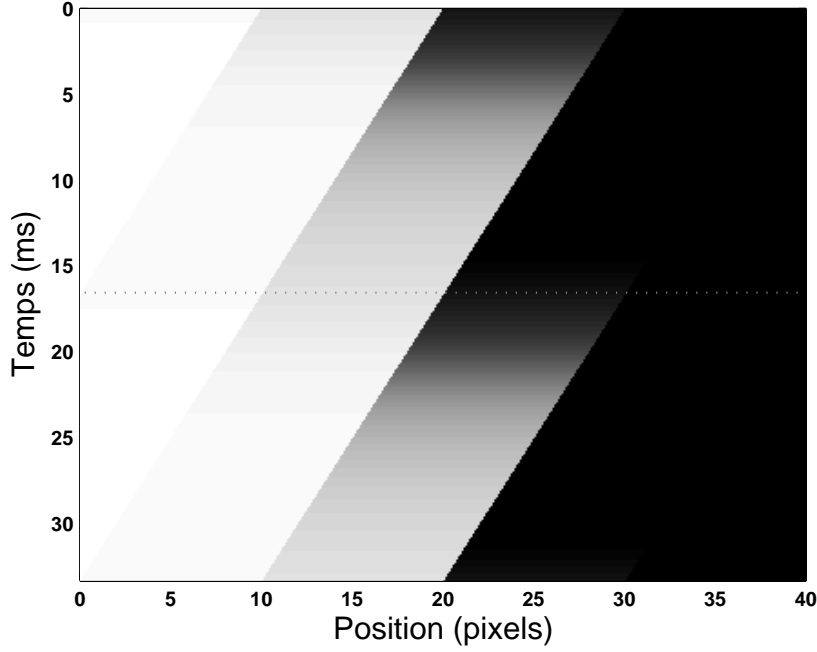


FIGURE 4.2 – Signal lumineux périodique du contour stabilisé sur le capteur $S(x, t)$, pour une durée de 2 images ($T = 16.67$ ms).

Le profil spatial du contour $\mathcal{C}_{N_i \rightarrow N_f}(x)$ est alors purement spatial et s'écrit :

$$\mathcal{C}_{N_i \rightarrow N_f}(x) = \lim_{\alpha \rightarrow +\infty} \frac{1}{\alpha} \int_{-\frac{\alpha}{2}}^{+\frac{\alpha}{2}} S(x, t) dt \quad (4.4)$$

Le signal incident $S(x, t)$ étant périodique de période T , on peut réduire les bornes de l'intégrale à n'importe quel intervalle de grandeur T :

$$\mathcal{C}_{N_i \rightarrow N_f}(x) = \int_0^T S(x, t) dt \quad (4.5)$$

$$\mathcal{C}_{N_i \rightarrow N_f}(x) = \int_0^T \mathcal{R}_{N_i \rightarrow N_f} \left(t - \left\lfloor \frac{x}{V} + \frac{t}{T} \right\rfloor T \right) dt \quad (4.6)$$

En choisissant l'intervalle $[-\frac{x}{V}T, T(1 - \frac{x}{V})]$, sur lequel l'élément $\lfloor \frac{x}{V} + \frac{t}{T} \rfloor$ est nul, on obtient :

$$\mathcal{C}_{N_i \rightarrow N_f}(x) = \int_{-\frac{x}{V}T}^{T(1 - \frac{x}{V})} \mathcal{R}_{N_i \rightarrow N_f}(t) dt \quad (4.7)$$

L'intégrale peut alors être étendue sur un intervalle infini en multipliant la transition temporelle par une fonction porte Π de largeur T et d'amplitude 1, centrée sur l'intervalle considéré :

$$\mathcal{C}_{N_i \rightarrow N_f}(x) = \int_{-\infty}^{+\infty} \mathcal{R}_{N_i \rightarrow N_f}(t) \cdot \Pi_T(t - \frac{x}{V}T) dt \quad (4.8)$$

Dans cette dernière expression, on reconnaît l'expression classique d'une corrélation. La fonction porte Π étant paire on peut réécrire cette expression comme une convolution. On obtient donc :

$$\begin{aligned}\mathcal{C}_{N_i \rightarrow N_j}(x) &= \mathcal{F}\left(\frac{x}{V}T\right) \\ \text{avec } \mathcal{F}(t) &= (\mathcal{R}_{N_i \rightarrow N_f} * \Pi_T)(t)\end{aligned}\tag{4.9}$$

4.2.4 Conséquence

L'analyse de la formation du flou de mouvement permet de montrer que le profil spatial d'un contour en mouvement $\mathcal{C}_{N_i \rightarrow N_j}(x)$ peut être obtenu en effectuant la convolution de la transition temporelle $\mathcal{R}_{N_i \rightarrow N_j}(t)$ par une fonction porte de largeur égale à la période de rafraîchissement de l'écran T . Ces deux fonctions sont converties du domaine temporel au domaine spatial par une mise à l'échelle entre la période de rafraîchissement T et la vitesse de déplacement du contour V ($V \equiv T$).

Cette méthode ne nécessite donc plus l'utilisation d'un dispositif imageur (caméra) puisque la connaissance de la réponse temporelle du pixel à une transition suffit pour en déduire le profil spatial d'un contour en mouvement. De plus, il n'est pas nécessaire d'utiliser un stimulus en mouvement comme pour la méthode précédente.

4.3 Mesures du flou de mouvement

Afin d'évaluer cette méthode analytique, une campagne de mesures du flou de mouvement a été menée avec les deux protocoles suivants :

- le protocole de mesure décrit dans le chapitre précédent, utilisant une caméra rapide pour obtenir l'image d'un contour en mouvement ;
- la méthode analytique développée ci-dessus, permettant l'obtention du profil spatial d'un contour en mouvement à partir de la réponse indicielle à la transition considérée.

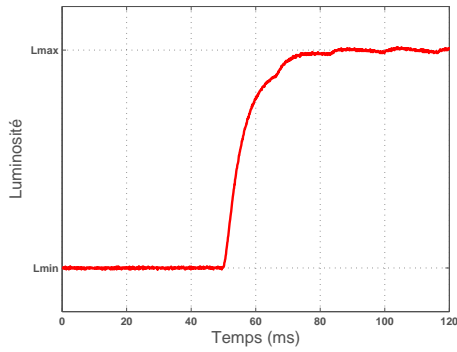
Cette campagne de mesure a été mise en œuvre sur quatre moniteurs informatiques, de manière à comparer les résultats obtenus.

4.3.1 Écrans testés

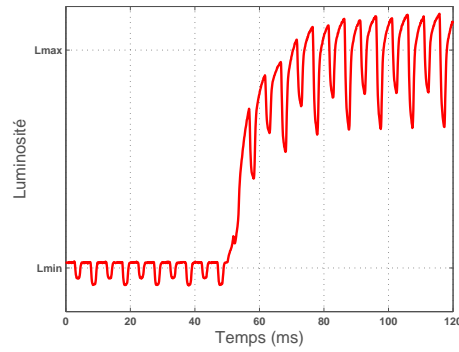
Le flou de mouvement a été mesuré sur quatre moniteurs informatiques LCD récents. Leurs caractéristiques sont données dans le tableau 4.1. Tous ces écrans à cristaux liquides sont des TFT à matrice active, fonctionnant avec une fréquence d'affichage de 60 Hz. Dans la suite, ils sont identifiés par les lettres A, B, C et D. Les écrans C et D sont équipés d'un système de clignotement du rétro-éclairage (*backlight flashing*, BF) pour réduire le flou de mouvement (cf. section 3.2.3).

Id	Type	Diagonale	Résolution	L_{max} (cd/m ²)	τ_{spec} (ms)	BF
A	S-IPS	20"	1600 × 1200	300	8	
B	S-PVA	24"	1920 × 1200	400	5	
C	S-IPS	26"	1920 × 1200	500	5	oui
D	S-IPS	30"	2560 × 1600	370	5	oui

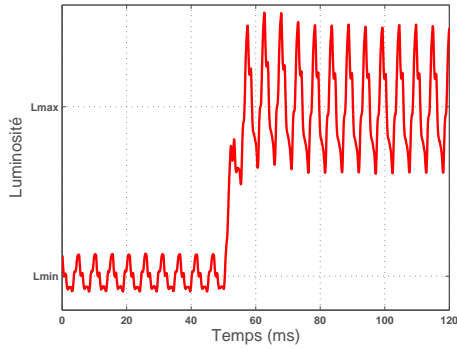
TABLE 4.1 – Spécifications des écrans mesurés. L_{max} est la luminance du blanc, τ_{spec} est le temps de réponse indiqué par le constructeur. BF indique que l'écran est équipé d'un système de clignotement du rétro-éclairage (*backlight flashing*) pour réduire le flou de mouvement (cf. section 3.2.3).



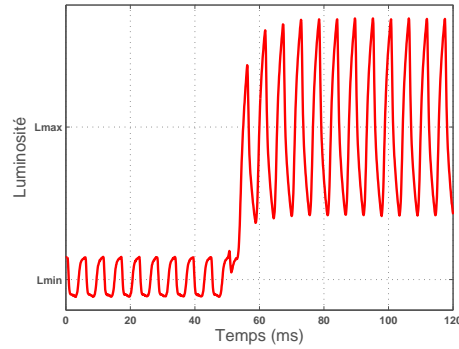
(a) Écran A.



(b) Écran B.



(c) Écran C.



(d) Écran D.

FIGURE 4.3 – Réponse indicielle temporelle des quatre écrans testés, pour la transition $0 \rightarrow 255$.

4.3.2 Mesures

4.3.2.1 À partir de la réponse indicielle

Le flou de mouvement a été mesuré sur 20 contours impliquant les 20 transitions gris-à-gris $N_i \rightarrow N_f$, telles que $(N_i, N_f) \in \{0, 63, 127, 191, 255\}^2$ et $N_i \neq N_f$.

Pour chaque transition $N_i \rightarrow N_f$, la réponse indicielle a été mesurée 5 fois, selon le protocole de mesure décrit section 3.3.1. La moyenne des 5 réalisations permet de réduire le bruit d'acquisition. La figure 4.3 illustre les réponses indiciaires obtenues pour les 4 écrans testés, pour la transition du noir au blanc $0 \rightarrow 255$. On peut remarquer le clignotement du rétro-éclairage sur les écrans C et D, ainsi que la modulation de largeur d'impulsions sur l'écran B.

Pour mesurer le temps de réponse des transitions mesurées (cf. section 3.6), ces réponses indiciaires ont été filtrées avec un filtre réjecteur de bande, de manière à s'affranchir des fréquences parasites dues au clignotement ou à la modulation de largeur d'impulsion du rétro-éclairage. Le temps de réponse τ a ensuite été mesuré sur le signal filtré, conformément aux recommandations [VESA 05].

Le profil spatial du contour flou $\mathcal{C}_{N_i \rightarrow N_f}(x)$ est obtenu pour chaque transition et pour chaque écran en utilisant la méthode analytique décrite section 4.2. Plus particulièrement, la convolution énoncée dans l'équation 4.9 est appliquée sur la réponse indicielle $\mathcal{R}_{N_i \rightarrow N_f}(t)$. Cette opération est effectuée sur la réponse indicielle brute, sans qu'il soit nécessaire d'effectuer le moindre filtrage (contrairement à ce qui est fait pour la mesure du temps de réponse). En effet, les modulations de luminance, et particulièrement le clignotement du rétro-éclairage, jouent un rôle important sur la largeur finale du flou de mouvement et doivent être prises en compte.

La figure 4.4 présente les profils spatiaux obtenus à partir des réponses indiciaires de la figure 4.3. Les données obtenues sont fonction du temps, le temps de flou $BET_{N_i \rightarrow N_f}$ peut donc être directement mesuré de la façon décrite section 3.4.3. Cependant, on peut se ramener à une dimension spatiale en prenant pour facteur d'échelle la vitesse V de déplacement du contour pour laquelle on souhaite prédire la largeur de flou. La largeur de flou $BEW_{N_i \rightarrow N_f}$ est donc obtenue en multipliant $BET_{N_i \rightarrow N_f}$ par V .

4.3.2.2 Avec un dispositif imageur

Le flou de mouvement a également été mesuré selon la méthode classique utilisant une caméra rapide, en suivant le protocole de mesure décrit section 3.4.2. Les mesures ont été effectuées sur 20 contours différents, impliquant les 20 transitions gris-à-gris $N_i \rightarrow N_f$, telles que $(N_i, N_f) \in \{0, 63, 127, 191, 255\}^2$ et $N_i \neq N_f$. Le contour capturé par la caméra était animé d'une vitesse de translation $V = 10$ pixels par image.

Les profils spatiaux obtenus pour les quatre écrans, pour un contour en mouvement impliquant la transition $0 \rightarrow 255$, sont donnés figure 4.5. La largeur de flou $BEW_{N_i \rightarrow N_f}$ est mesurée sur le profil obtenu, le temps de flou $BET_{N_i \rightarrow N_f}$ est calculé par la relation 3.15, c'est à dire dans le cas présent :

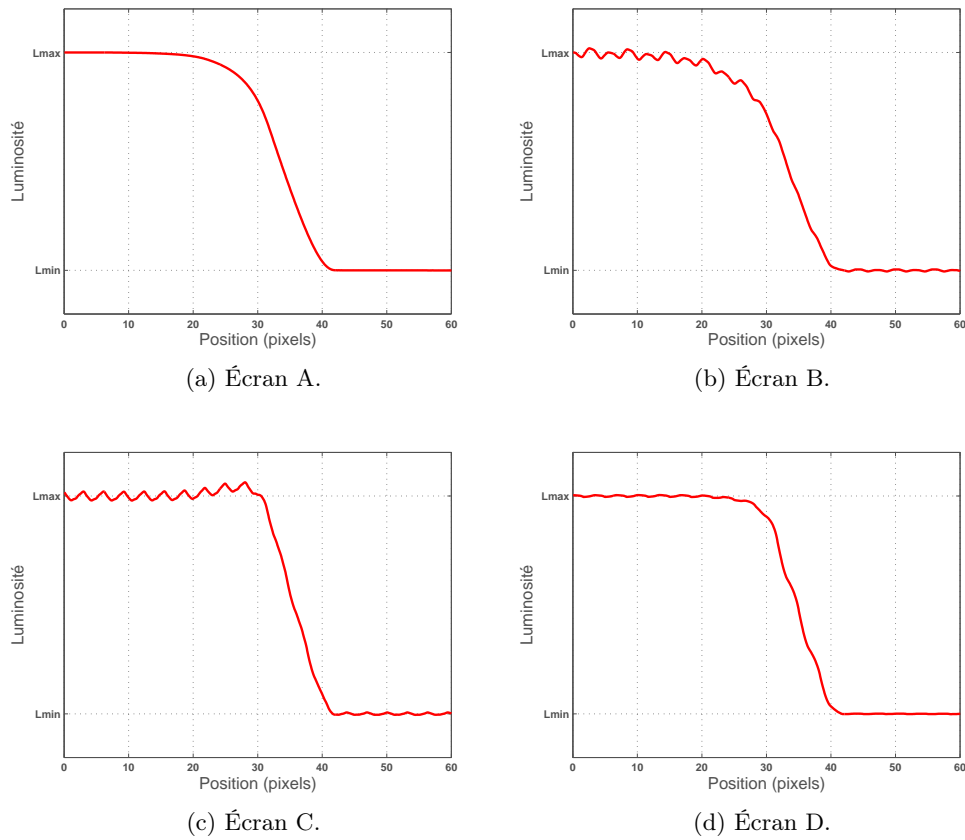


FIGURE 4.4 – Profils spatiaux d'un contour en mouvement obtenus à partir des réponses indicielles de la figure 4.3, et simulés pour une vitesse $V = 10$ pixels par image. La transition considérée est la transition $0 \rightarrow 255$.

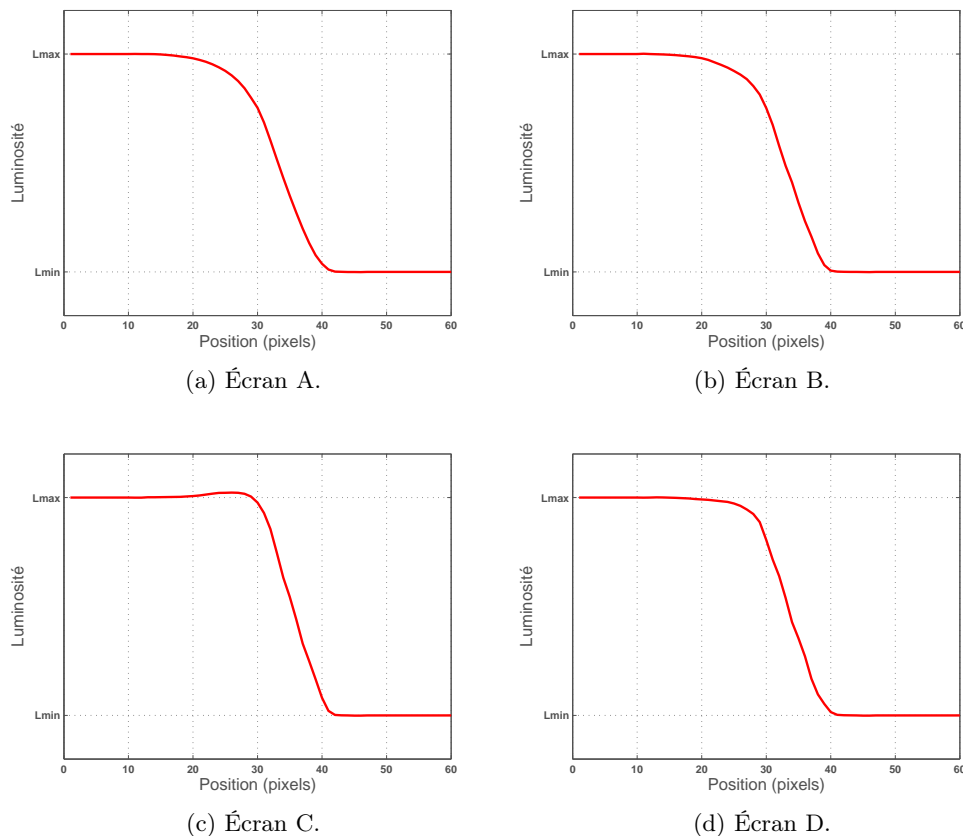


FIGURE 4.5 – Profils spatiaux d'un contour en mouvement, animé d'une vitesse $V = 10$ pixels par image, obtenus à partir des images capturées par la caméra rapide. La transition considérée est la transition $0 \rightarrow 255$.

$$BET_{N_i \rightarrow N_f} = \frac{BEW_{N_i \rightarrow N_f}}{10} \quad (4.10)$$

4.4 Résultats

Les tableaux 4.2 et 4.3 présentent les résultats de la campagne de mesure sur les quatre écrans testés, pour les 20 transitions gris-à-gris. Le temps de flou a été mesuré cinq fois pour chaque transition et pour chaque écran, et les valeurs recensées dans les tableaux 4.2 et 4.3 sont la moyenne des ces cinq mesures. Le temps de flou mesuré sur le profil spatial du contour en mouvement capturé par la caméra est noté BET^S . Le temps de flou mesuré sur le profil spatial du contour en mouvement obtenu à partir de la réponse indicielle est noté BET^T . Le temps de réponse τ est également rapporté pour chaque transition. Les valeurs moyennes de ces trois mesures sont calculées. Comme attendu, le temps de flou est le plus faible pour les écrans C et D sur lesquels le rétro-éclairage est clignotant de manière à réduire la quantité de flou formé sur la rétine.

	Écran A			Écran B		
Transition	BET^S	BET^T	τ	BET^S	BET^T	τ
$0 \rightarrow 63$	<i>27.0</i>	<i>20.9</i>	16.0	29.0	27.7	23.9
$0 \rightarrow 127$	<i>24.5</i>	<i>20.8</i>	15.3	29.7	30.8	25.9
$0 \rightarrow 191$	23.2	21.2	16.2	29.2	30.0	25.9
$0 \rightarrow 255$	20.8	19.7	13.6	<i>20.0</i>	<i>24.4</i>	22.2
$63 \rightarrow 0$	14.7	15.2	9.5	13.3	14.1	13.9
$63 \rightarrow 127$	22.7	20.7	15.0	27.3	28.2	25.0
$63 \rightarrow 191$	<i>22.7</i>	<i>20.4</i>	14.4	27.5	28.7	26.1
$63 \rightarrow 255$	20.2	19.4	13.1	<i>17.7</i>	<i>22.2</i>	22.4
$127 \rightarrow 0$	15.3	15.2	9.6	13.5	13.9	14.9
$127 \rightarrow 63$	18.7	18.3	12.2	19.0	19.2	11.0
$127 \rightarrow 191$	21.0	20.0	13.6	27.3	28.5	26.8
$127 \rightarrow 255$	19.5	19.1	12.7	<i>17.3</i>	<i>21.7</i>	21.7
$191 \rightarrow 0$	15.5	15.4	9.8	13.3	14.1	15.0
$191 \rightarrow 63$	18.0	17.5	11.5	<i>17.2</i>	<i>19.2</i>	10.9
$191 \rightarrow 127$	20.3	18.7	12.6	<i>23.7</i>	<i>26.2</i>	24.7
$191 \rightarrow 255$	19.2	18.9	12.4	<i>17.5</i>	<i>22.6</i>	22.6
$255 \rightarrow 0$	15.7	15.5	9.9	13.5	14.1	15.0
$255 \rightarrow 63$	18.2	17.2	11.3	19.7	20.8	13.9
$255 \rightarrow 127$	19.8	18.5	12.4	24.8	26.1	22.7
$255 \rightarrow 191$	20.0	19.2	13.1	27.0	29.3	24.4
Moyenne	19.8	18.6	12.7	21.4	23.1	20.4

TABLE 4.2 – Résultats des mesures pour les écrans A et B. Toutes les valeurs sont en millisecondes. BET^S est le temps de flou mesuré sur la profil spatial d'un contour en mouvement capturé par une caméra rapide. BET^T est le temps de flou mesuré sur le profil de contour flou obtenu à partir de la réponse temporelle indicielle de l'écran. Les cellules en *italique* correspondent aux valeurs de temps de flou pour lesquelles la différence entre les deux méthodes est supérieure à 10%.

	Écran C			Écran D		
Transition	BET^S	BET^T	τ	BET^S	BET^T	τ
$0 \rightarrow 63$	13.0	14.2	18.1	13.7	14.1	22.2
$0 \rightarrow 127$	15.5	14.3	22.3	15.5	15.0	21.0
$0 \rightarrow 191$	14.3	15.2	21.8	13.8	13.0	20.9
$0 \rightarrow 255$	13.5	13.9	18.7	15.7	15.6	9.7
$63 \rightarrow 0$	15.2	15.7	17.4	14.7	14.3	9.5
$63 \rightarrow 127$	12.8	13.7	24.8	14.3	13.9	22.6
$63 \rightarrow 191$	13.5	14.1	21.4	14.2	14.0	20.8
$63 \rightarrow 255$	13.7	13.7	18.4	15.2	15.5	9.6
$127 \rightarrow 0$	15.8	15.6	9.4	15.8	15.2	9.5
$127 \rightarrow 63$	15.3	14.7	10.7	15.8	15.2	9.9
$127 \rightarrow 191$	12.3	13.0	24.4	12.7	13.1	22.4
$127 \rightarrow 255$	13.2	13.3	18.8	15.3	15.4	9.4
$191 \rightarrow 0$	15.5	15.7	10.0	15.2	15.4	9.7
$191 \rightarrow 63$	15.7	15.3	9.6	15.7	15.6	10.2
$191 \rightarrow 127$	13.8	14.2	21.2	15.3	14.9	17.7
$191 \rightarrow 255$	13.0	12.6	22.4	14.8	15.4	9.4
$255 \rightarrow 0$	16.3	16.1	10.7	15.2	15.7	10.0
$255 \rightarrow 63$	15.8	15.6	9.9	16.0	15.9	10.3
$255 \rightarrow 127$	14.8	14.7	19.1	15.2	15.3	10.2
$255 \rightarrow 191$	12.5	12.7	23.6	13.2	14.0	21.1
Moyenne	14.3	14.4	17.6	14.9	14.8	14.3

TABLE 4.3 – Idem que le tableau 4.2 mais pour les écrans C et D.

4.4.1 Observations

Des différences significatives sont observées entre les mesures de temps de flou obtenues à partir de chaque méthode. Le tableau 4.4 recense l'erreur absolue moyenne ainsi que le coefficient de corrélation linéaire entre les deux séries de mesures, pour chaque écran.

Ces différences sont particulièrement importantes sur l'écran B, avec une erreur absolue moyenne de 1.85 ms. Sur l'écran A, l'erreur moyenne est assez élevée mais elle est due principalement aux fortes différences observées sur les transitions avec de faibles luminances ($0 \rightarrow 63$ et $0 \rightarrow 127$ notamment, cf. tableau 4.2). Pour les autres transitions les valeurs sont assez proches. Enfin les deux méthodes de mesure donnent des résultats très proches pour les écrans C et D. Lorsqu'on prend en compte l'ensemble des mesures sur les quatre écrans, l'erreur moyenne absolue est de 1 ms, soit environ 5.7% de la valeur moyenne du temps de flou, ce qui reste acceptable.

Le coefficient de corrélation linéaire entre les deux séries de mesure est très élevé pour les écrans A et B. Il est un peu plus faible pour les écrans C et D, pour lesquels le temps de flou sont concentrés

	Écran A	Écran B	Écran C	Écran D	Global
Erreur absolue moyenne	1.31	1.85	0.47	0.39	1.00
Coefficient de corrélation linéaire	0.9400	0.9594	0.8941	0.8770	0.9426

TABLE 4.4 – Erreur absolue moyenne (en millisecondes) et coefficient de corrélation linéaire entre le temps de flou BET^T et le temps de flou BET^S pour chacun des quatre écrans testés, et pour l'ensemble des mesures.

sur un faible intervalle (entre 12.3 ms et 16.3 ms). Globalement cependant, la corrélation entre les deux séries de mesure (80 mesures distinctes) est très bonne.

Concernant la stabilité des mesures effectuées, l'écart-type obtenu sur les cinq occurrences de chaque mesure est de 0.6 en moyenne pour les mesures à partir de la réponse indicielle, et de 0.4 en moyenne pour les mesures à partir des images acquises par caméra.

4.4.2 Comparaison des résultats obtenus par chacune des deux méthodes

La figure 4.6 compare les profils spatiaux des contours en mouvement obtenus avec les deux méthodes. Pour chaque écran, la transition choisie est celle pour laquelle l'erreur absolue est la plus forte. Plusieurs raisons peuvent expliquer ces variations.

Tout d'abord, des artefacts temporels apparaissent sur les profils spatiaux des écrans B et C obtenus par la convolution de la réponse indicielle avec une fenêtre de largeur égale à la période d'affichage de l'écran. Ces artefacts sont les résidus des modulations temporelles présentes sur les réponses indicielles de ces deux écrans. Ces modulations sont dues à la modulation de largeur d'impulsion effectuée pour régler la luminance du rétro-éclairage sur l'écran B, et au clignotement du rétro-éclairage en vue de réduire le flou de mouvement sur l'écran C. Elles ne sont pas annulées par la convolution parce que leur fréquence n'est pas un multiple de la fréquence d'affichage de l'écran (la fréquence de la MLI est de 204 Hz sur l'écran B et la fréquence du clignotement est de 192 Hz sur l'écran C). En effet, la convolution avec une fenêtre de largeur égale à la période d'affichage revient à multiplier le spectre de la réponse indicielle par une fonction sinus cardinal dont les passages à zéro ont lieu pour tous les multiples non-nuls de la fréquence d'affichage de l'écran. Pour cette raison, les modulations temporelles de l'écran D ne sont pas observées sur le profil spatial obtenu par cette méthode, car la fréquence du clignotement du rétro-éclairage sur cet écran est égale à 180 Hz, soit 3 fois la fréquence d'affichage. Sur les écrans B et C, ces modulations sont atténuées mais pas totalement supprimées.

Ce résultat est important car il n'avait pas été relevé sur les récentes études présentant le même type de protocole. En effet, aussi bien dans les travaux de Becker [Becker 08b] que dans ceux de Feng *et al.* [Feng 08], l'écran testé présentait des artefacts temporels dus à une MLI de fréquence égale à un multiple de la fréquence d'affichage (225Hz pour 75 Hz dans le premier cas, 120 Hz pour 60 Hz dans le second cas). Il est cependant important d'être conscient que dans le cas où les modulations temporelles (MLI ou clignotement du rétro-éclairage) n'ont pas une fréquence égale à un multiple

de la fréquence d'affichage, des résidus seront présents sur le profil spatial obtenus à partir de la réponse indicielle. L'amplitude de ces résidus n'est pas très élevée dans notre cas, mais ils peuvent potentiellement affecter la mesure du temps de flou. Il peut donc être nécessaire de les filtrer à l'aide d'un filtre réjecteur de bande réglé sur la fréquence incriminée, en veillant à ce que cette opération n'altère pas la largeur du contour.

A l'inverse, les résultats obtenus par la méthode classique de capture du contour en mouvement par une caméra rapide sont insensibles à tout artefact temporel. En effet, l'intégration temporelle du capteur pour chaque image acquise est plus longue, et la fréquence d'échantillonnage temporelle beaucoup plus faible par rapport à celle de la photo-diode utilisée pour mesurer la réponse indicielle. Par conséquent, les hautes fréquences sont directement filtrées par la caméra. De plus, la sommation temporelle des images acquises participe également au lissage des contours obtenus.

On peut aussi remarquer que les différences observées entre les deux méthodes peuvent être due à des artefacts liés à la mesure par caméra rapide. Sur l'écran A par exemple (cf. figure 4.6a), la différence entre les deux temps de flou mesuré est clairement due à une mauvaise estimation du contour à partir des images capturées. Cela se traduit également par un écart-type élevé entre les cinq mesures effectuées (1.85 pour $0 \rightarrow 63$, 1.36 pour $63 \rightarrow 0$). En effet, pour les transitions de faible luminance (particulièrement $0 \rightarrow 63$ et $0 \rightarrow 127$) les images capturées par la caméra peuvent être assez bruitées. De plus la faible dynamique de luminance (particulièrement pour l'écran A qui est celui avec la plus faible luminance maximale parmi les écrans testés) du contour peut accentuer les effets du bruit d'acquisition.

Au final, en considérant les résultats de mesures obtenus avec les deux protocoles, les constatations suivantes peuvent être faites.

Concernant la mesure du flou de mouvement à partir de la réponse indicielle :

- La fréquence d'échantillonnage plus élevée permet d'obtenir des résultats plus précis.
- Ce protocole ne nécessite aucun traitement d'image, ni compensation de mouvement.
- La mesure est facile à mettre en œuvre, et peut être reproductible d'un laboratoire à l'autre avec des conditions identiques.
- Cependant, dans le cas où des modulations temporelles sont présentes sur la réponse indicielle (clignotement ou modulation par largeur d'impulsions du rétro-éclairage par exemple) il peut être nécessaire de filtrer les résidus temporels, dans le cas seulement où la fréquence de ces modulations n'est pas un multiple de la fréquence d'affichage de l'écran. Ce traitement peut introduire un biais sur les résultats obtenus et sur la reproductibilité d'un laboratoire à l'autre.

Concernant la mesure du flou de mouvement à partir des images du contour acquises par une caméra rapide :

- Ce protocole nécessite un dispositif matériel plus important, complexe et coûteux.
- Le traitement des données brutes, le type de matériel utilisé, le calibrage de l'ensemble peuvent conduire à de fortes différences de résultats d'un laboratoire à l'autre.
- Les résultats peuvent être sensibles au bruit d'acquisition, particulièrement pour les transitions de faible luminance. Les résultats sont alors moins stables et l'écart entre deux mesures peut-être

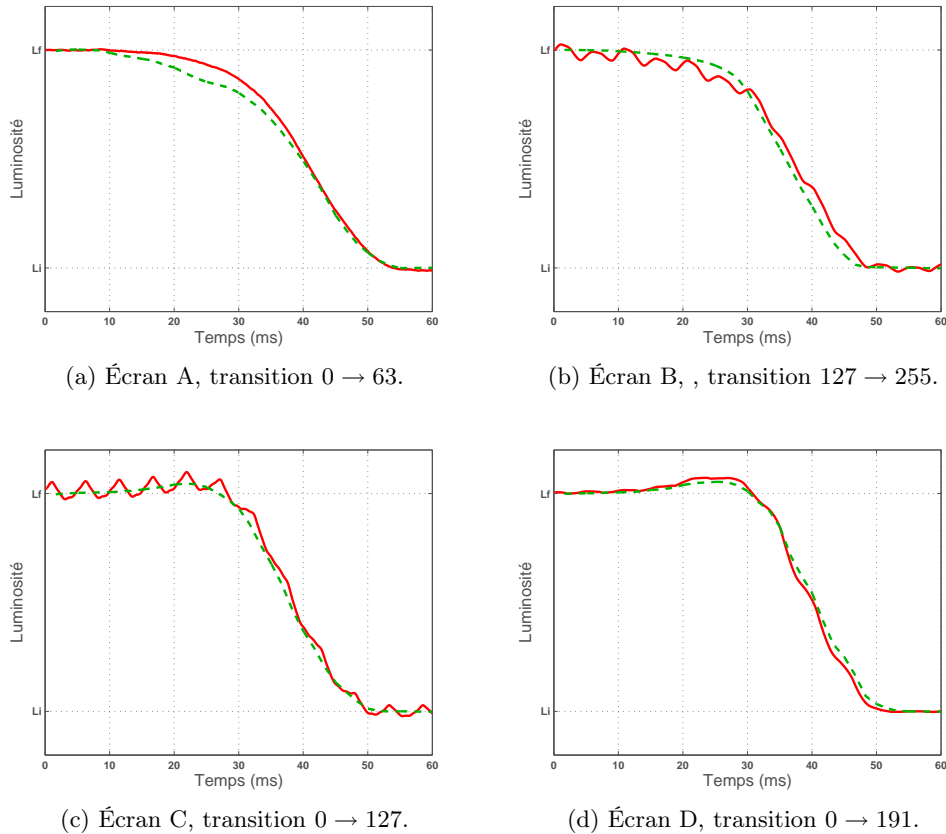


FIGURE 4.6 – Comparaison entre les profils spatiaux d'un contour en mouvement obtenus avec les deux méthodes de mesure. Pour chaque écran, la transition choisie est celle pour laquelle l'écart entre les mesures est le plus important. Les courbes pleines sont obtenues à partir des réponses indicielles, les courbes pointillées sont obtenues à partir des images capturées par la caméra rapide.

assez élevé.

- Cependant, ce type de mesure est moins sensible aux modulations temporelles du rétro-éclairage et permet d'obtenir un profil spatial du contour en mouvement vierge de tout résidu.

4.4.3 Relation entre temps de réponse et temps de flou

On peut également profiter des résultats obtenus par cette campagne de mesures pour étudier la relation entre le temps de réponse et le temps de flou. Dans un premier temps, on remarque que les valeurs mesurées pour le temps de réponse sont bien loin de celles données par les constructeurs dans le tableau 4.1. Cela confirme si besoin est la nécessité de définir une mesure standard du temps de réponse de manière à pouvoir comparer les valeurs d'un écran à l'autre, ce qui est actuellement impossible avec les valeurs données dans les spécifications techniques des écrans du commerce.

Cependant, la question principale est de savoir quelle est la légitimité du temps de réponse pour caractériser le comportement temporel des écrans à cristaux liquides. Le tableau 4.5 présente la corrélation entre le temps de réponse et le temps de flou pour chacun des 4 écrans testés. Pour l'écran

	Écran A	Écran B	Écran C	Écran D	Global
$\text{corr}(BET^T, \tau)$	0.9776	0.8658	-0.7628	-0.8180	0.4015
$\text{corr}(BET^S, \tau)$	0.9710	0.7850	-0.8072	-0.7441	0.2873

TABLE 4.5 – Coefficient de corrélation linéaire entre les temps de réponse τ et les temps de flou BET^T pour chacun des quatre écrans testés, et pour l'ensemble des mesures.

A, dont les réponses indicielles sont vierges de tout artefact temporel, les deux séries de mesures sont hautement corrélées. Le temps de réponse peut donc être considéré comme un indicateur fiable de la quantité de flou de mouvement formé sur l'écran. En d'autres termes, plus le temps de réponse est élevé plus le flou de mouvement sera important.

Pour les écrans dont la réponse indiciale présente des artefacts temporels, les valeurs sont bien moins corrélées. Dans le cas des écrans C et D, le clignotement du rétro-éclairage est mis en œuvre pour réduire la formation du flou de mouvement sur la rétine. Dans ce contexte la mesure du temps de réponse n'a plus de sens et le besoin d'un nouvel indice pour caractériser le comportement temporel de l'écran est nécessaire, particulièrement pour les écrans récents ayant des réponses indicielles « impulsives ». Le temps de flou est même inversement proportionnel au temps de réponse pour les écrans C et D : les transitions avec le temps de réponse le plus faible sont celles qui obtiennent un flou de mouvement le plus élevé.

On remarque par exemple que le temps de réponse moyen de l'écran A est inférieur à celui de l'écran C de 28%, alors que le temps de flou moyen mesuré sur l'écran A est supérieur de 28% à celui mesuré sur l'écran C. La prise en compte du temps de réponse pour comparer deux écrans peut donc amener à de mauvaises conclusions en termes de quantité de flou de mouvement formé sur la rétine. Cela est confirmé par la corrélation globale entre les deux valeurs. Cette mesure n'est donc plus valable quand il s'agit de caractériser les écrans équipés de systèmes de réduction du flou de mouvement.

4.5 Conclusion

Dans ce chapitre nous avons mis en œuvre deux protocoles de mesure du flou de mouvement. Le premier consiste à acquérir un contour en mouvement avec une caméra rapide et à associer les images obtenues de manière à obtenir une image du contour flou. Le second protocole permet l'obtention du profil spatial du contour flou directement à partir de la réponse indiciale de l'écran pour la transition considérée.

Le flou de mouvement a été mesuré sur quatre écrans avec des réponses indicielles très différentes, et pour 20 transitions gris-à-gris. Les mesures obtenues nous permettent de confirmer que le flou de mouvement peut être mesuré à partir de la réponse temporelle de l'écran comme cela avait été simulé dans différentes études [Pan 05, Watson 06]. En effet, les résultats obtenus via les deux protocoles sont très proches avec une erreur moyenne absolue de l'ordre de 1 ms et un coefficient de corrélation

linéaire supérieur à 0.94. Deux études très récentes [Becker 08b, Feng 08] ont également montré la correspondance entre ces deux approches. Cependant, nous avons montré que, sous certaines conditions, des différences pouvaient apparaître entre les deux mesures, particulièrement lorsque des modulations temporelles du rétro-éclairage sont présentes. Ce phénomène peut être corrigé lorsqu'il se présente.

La méthode de mesure consistant à obtenir le contour flou à partir de la réponse indicielle permet une mesure fiable et rapide des écrans. De plus, elle est facilement reproductible d'un laboratoire à l'autre. Pour ces raisons, elle a été choisie par TCO Development consécutivement à ce projet de recherche, pour les futures mesures de flou de mouvement requises dans la version à venir du label TCO Media Displays. Elle devrait donc être déployée prochainement dans les laboratoires de tests de l'organisme de certification.

Néanmoins, la mesure du flou de mouvement n'est qu'une étape dans la compréhension du flou de mouvement. Elle permet de connaître le profil spatial d'un contour en mouvement tel qu'il parvient sur la rétine (si les effets de l'optique de l'œil sont négligés). Cependant, pour définir des recommandations concernant les caractéristiques temporelles d'un écran, il faut étudier plus précisément la *perception* du flou de mouvement. C'est l'objet de la partie suivante, où la réponse du système visuel humain au flou de mouvement est étudiée à travers un ensemble de tests psychophysiques.

Troisième partie

Étude psychophysique du flou de mouvement

Chapitre 5

Influence du mouvement sur la réponse supraliminaire du système visuel humain

Sommaire

5.1	Introduction	83
5.2	Contexte	84
5.2.1	Fonctions de sensibilité au contraste	84
5.2.2	Fonctions de sensibilité au contraste en mouvement	85
5.2.3	Fonctions de sensibilité supraliminaires	86
5.3	Expérimentations psychophysiques	87
5.3.1	Observateurs	87
5.3.2	Matériel	87
5.3.3	Stimuli et conditions de visualisation	87
5.3.4	Procédure	89
5.4	Résultats	90
5.4.1	Données mesurées	90
5.4.2	Modélisation du gain	92
5.4.3	Surface spatio-temporelle du gain en sensibilité	94
5.5	Discussion	95
5.5.1	Stabilisation du contraste	95
5.5.2	Comparaison avec les mesures effectuées au seuil de visibilité	98
5.6	Conclusion	100

5.1 Introduction

La deuxième partie de ce mémoire a abordé la mesure physique et objective du flou de mouvement. Nous y avons décrit des techniques métrologiques permettant d'obtenir le profil spatial d'un contour en mouvement. Celui-ci est mesuré à partir des images acquises par une caméra rapide, ou bien directement à partir de la réponse indicielle temporelle de l'écran. Si on considère que les distorsions

apportées par l'optique de l'œil sont négligeables, le profil spatial obtenu correspond à l'image du contour affichée sur la rétine lorsque l'observateur suit des yeux un objet en mouvement sur un écran à cristaux liquides (mouvement oculaire de poursuite). Cependant, de nombreux processus neuronaux interviennent au niveau de la rétine ainsi que dans le cortex visuel. Cela implique que le contour en mouvement perçu par le système visuel humain diffère du profil spatial obtenu par des mesures purement optiques. Dans cette troisième partie, nous nous intéressons donc plus particulièrement à la perception du flou de mouvement. Cette étude doit nous permettre de comprendre et de quantifier la réponse du système visuel humain à ce stimulus particulier. Par l'intermédiaire d'expériences psychovisuelles, nous allons mesurer cette perception et comparer les données subjectives obtenues avec les mesures physiques effectuées précédemment.

Dans ce chapitre, nous avons décidé d'étudier la réponse du système visuel humain à un stimulus sinusoïdal en mouvement sur la rétine, de manière à explorer l'influence du mouvement sur la perception. Le SVH est la combinaison d'une multitude de processus, qui modifient le signal lumineux capté par l'œil en un signal pouvant être interprété par notre cerveau. L'une des approches les plus classiques pour tenter de mesurer la sensibilité du SVH à un certain type de stimulus consiste à mesurer le seuil de visibilité du stimulus. Ce genre de mesure permet d'obtenir des fonctions de sensibilité au contraste (CSF) qui donnent le contraste minimum nécessaire pour qu'un observateur puisse détecter la présence d'un stimulus, bien entendu ce contraste seuil dépend de nombreux paramètres tels que les fréquences spatiales et temporelles du stimulus, mais aussi les conditions d'observation, l'éclairage ambiant, la luminosité moyenne du stimulus, etc. (voir Annexe A). Puisque nous nous intéressons ici à un stimulus précis, d'une magnitude généralement bien au delà du seuil différentiel de visibilité, nous avons décidé de mesurer la réponse supraliminaire (au delà du seuil de perception) du SVH à des stimuli en mouvement.

5.2 Contexte

5.2.1 Fonctions de sensibilité au contraste

La réponse du système visuel humain au seuil de visibilité a été largement étudiée, depuis les travaux initiaux de Schade [Schade 56], de Lowry & DePalma [Lowry 61, DePalma 62] et de Campbell & Green [Campbell 65], jusqu'aux modèles exhaustifs de Barten [Barten 99] plus récemment. Les fonctions de sensibilité aux contrastes (CSF) mesurées sur la composante de luminance ont typiquement un comportement passe-bande : la sensibilité est maximale entre 3 et 5 cycles par degré (cpd) et décroît aussi bien pour les hautes fréquences que pour les basses fréquences (cf. annexe A.5.2). Ce comportement peut être expliqué par l'interaction de deux mécanismes basiques du système visuel humain [DePalma 62] :

- un mécanisme passe-bas lié à la qualité de l'optique de l'œil et aux limites de résolution du SVH [Campbell 65], il est responsable de la chute de sensibilité aux hautes fréquences spatiales ;
- un mécanisme passe-haut lié aux processus inhibiteurs qui interviennent entre la rétine et le cortex visuel.

Les fonctions de sensibilité au contraste peuvent également être mesurées en fonction de caractéristiques temporelles du stimulus. La fréquence temporelle [Robson 66], ainsi que la durée de présentation du stimulus [Nachmias 67], influence la sensibilité de l'observateur. Les courbes typiques des fonctions de sensibilité aux fréquences temporelles sont données dans l'annexe A (section A.5.2.2), leur allure est similaire à celle des fonctions de sensibilité aux fréquences spatiales. De plus, les CSF spatiales et temporelles sont séparables pour des hautes fréquences spatiales et temporelles.

5.2.2 Fonctions de sensibilité au contraste en mouvement

Dans la nature, les objets d'intérêt sont généralement animés de mouvement plutôt qu'en état de clignotement. En conséquence, les variations temporelles rencontrées par l'œil seront plus probablement dues à des stimuli spatiaux en mouvement sur la rétine. Pour cette raison, il est également intéressant de mesurer la sensibilité du système visuel humain à des stimuli en mouvement. La CSF obtenue est une forme particulière de CSF spatio-temporelle, elle est désignée dans la littérature par l'expression *spatio-velocity CSF*. Nous utiliserons dans ce mémoire le terme de CSF en mouvement. L'expression mathématique d'un motif de Gabor animé d'un mouvement de translation horizontale peut être déduite de la forme spatiale du motif de Gabor décrit en annexe (équation A.9) :

$$\text{Gabor}(x, y, t) = \text{Gabor}(x + Vt, y) \quad (5.1)$$

avec V la vitesse de déplacement du stimulus.

Le stimulus traverse alors l'écran de la gauche vers la droite, animé d'une vitesse V en degrés par seconde. Puisqu'on cherche à mesurer la sensibilité au contraste pour des fluctuations temporelles sur la rétine, il faut que celle-ci reste stationnaire. En effet, si l'observateur suit des yeux le mouvement du stimulus ce dernier sera stabilisé sur la rétine et aucune variation temporelle n'aura lieu au niveau des photo-récepteurs. Laird *et al.* ont montré que la CSF mesurée dans ces conditions était similaire à une CSF spatiale [Laird 06]. Un autre problème réside dans le fait que si le stimulus traverse l'écran et que la rétine reste stationnaire (fixée sur un point au centre de l'écran par exemple), les zones périphériques de la rétine seront excitées. Étant donnée l'inhomogénéité de la rétine, la réponse du système visuel humain à ce stimulus tiendra alors compte de la réponse extra-fovéale, différente. Une solution consiste à modifier le stimulus de manière à ce que la fenêtre gaussienne reste stationnaire au centre de l'écran, seul la sinusoïde étant animée de mouvement :

$$\text{Gabor}(x, y, t) = A \left[\sin(2\pi f_s(x + Vt)) \cdot \exp\left(-\frac{x^2 + y^2}{0.36 \cdot l^2}\right) \right] + L_{moy} \quad (5.2)$$

Pour les deux types de stimuli précédemment décrits, la fréquence des fluctuations temporelles en un point donné de la rétine est la suivante :

$$f_t = f_s \cdot V \quad (5.3)$$

Parmi les travaux d'importance dans ce domaine, on peut notamment citer ceux de Koenderink [Koenderink 78, Koenderink 79] et ceux de Kelly [Kelly 79b, Burbeck 80, Kelly 84, Kelly 85]. Ce

dernier a montré que la sensibilité au contraste pour des stimuli en mouvement était plus grande que pour des stimuli clignotants (avec un facteur de l'ordre de 2). Il a développé un modèle de CSF en deux dimensions qui a été par la suite amélioré par Daly [Daly 98], grâce à la prise en compte des différents types de mouvements de l'œil ainsi que ses limitations physiques en termes de vitesse.

5.2.3 Fonctions de sensibilité supraliminaires

Il est commun d'identifier la CSF au seuil de visibilité comme la fonction de transfert de modulation du système visuel. Cependant, Lowry et DePalma [Lowry 61] ont montré que la CSF ne pouvait être considérée comme la réponse du système visuel humain à des niveaux supraliminaires (très au delà du seuil différentiel de visibilité). Durant les années 1960 et 1970, de nombreux travaux ont été conduits [Bryngdahl 66, Watanabe 68, Georgeson 75, Kulikowski 76] pour mesurer la fonction de sensibilité au contraste supraliminaire, c'est à dire la variation du contraste perçu au delà du seuil de visibilité, en fonction de la fréquence spatiale. Toutes ces études ont montré que la réponse du système visuel à des niveaux supraliminaires ne pouvait être dérivée linéairement de la CSF au seuil de visibilité. Plus précisément, il apparaît que la variation du contraste perçu avec la fréquence spatiale devient de moins en moins prononcée au fur et à mesure que le contraste du stimulus augmente. Une hypothèse serait que le mécanisme de détection au seuil de visibilité est différent du mécanisme de perception au delà du seuil. Cette hypothèse est vérifiée par les travaux de Medjeur & Tulunay-Keeseey [Medjbeur 86] impliquant des observateurs possédant de fortes anomalies visuelles au seuil de visibilité. Pour une majorité de ces sujets, la réponse du système visuel humain pour de forts contrastes est semblable à celle de sujets sains. De plus, des auteurs, tels que Cannon [Cannon 85] et Legge & Kersten [Legge 87], ont montré que la perception du contraste à des niveaux supraliminaires ne variait pas avec l'excentricité, tandis que cette dernière joue un rôle très important dans la détermination du seuil différentiel de visibilité.

De la même façon, la réponse du système visuel humain à des niveaux de contraste supraliminaires dans le domaine temporel ne peut être simplement déduite des fonctions de sensibilité au contraste mesurées au seuil de visibilité [Georgeson 87, Waugh 94, Manahilov 03, Bex 07]. La réponse du système visuel humain a été mesurée dans le domaine spatio-temporel par Bowker [Bowker 83] en utilisant des sinusoïdes clignotantes (cf. équation A.11). Une procédure d'égalisation de contraste entre un stimulus constant et le stimulus test a été utilisée. Cette étude montre que la réponse supraliminaire du système visuel humain possède un comportement similaire dans le domaine temporel à celui observé dans un domaine purement spatial. L'influence du contenu fréquentiel, aussi bien temporel que spatial, du stimulus est de moins en moins importante à mesure que son contraste augmente. Cette étude des stimuli spatio-temporels à des niveaux de contrastes supraliminaires n'a cependant pas été menée pour des stimuli en mouvement.

Dans le cadre de ces travaux de thèse, nous nous intéressons plus particulièrement aux mécanismes impliqués dans la perception du flou de mouvement. Or comme nous l'avons montré dans le chapitre 4, la formation du flou de mouvement sur la rétine est provoquée par le déplacement d'un stimulus dans le référentiel rétinien. Les cellules rétiniennes sont alors soumises à un signal incident comportant une variation temporelle avec continuité spatiale (cf. figure 4.2). Pour tenter de comprendre la

réponse du système visuel humain à ce type d'excitation, nous avons donc décidé de mesurer celle-ci à des niveaux de contraste supraliminaire, et pour des stimuli en mouvement.

5.3 Expérimentations psychophysiques

Pour mesurer l'influence du mouvement sur la réponse du système visuel humain à des niveaux supraliminaire, nous avons développé une expérience psychophysique d'égalisation de contraste. L'égalisation de contraste est faite entre deux motifs sinusoïdaux, l'un en mouvement, l'autre stationnaire. Les deux stimuli possèdent la même fréquence spatiale.

5.3.1 Observateurs

Cinq observateurs ont participé à l'expérience. Tous possèdent une vision normale (acuité visuelle de 9/10 minimum), avec correction le cas échéant. Ils ont été familiarisés avec la procédure après plusieurs sessions d'entraînement dont les résultats n'ont pas été pris en compte.

5.3.2 Matériel

Les stimuli ont été présentés sur un écran à tube cathodique Samsung SyncMaster 1100MB d'une diagonale de 21 pouces (53.3 cm). La luminance moyenne de l'écran a été fixée à 50 cd/m², sa résolution est de 1024 × 768 pixels et sa fréquence d'affichage de 120 Hz. La luminance de l'écran a été mesurée pour chaque valeur de chacune des composantes RGB avec un photomètre OptiCAL OP200-E (Cambridge Research System). Cette mesure a permis de générer des stimuli dont le contraste est linéaire en luminance, avec une résolution de 8 bits. L'équilibre des couleurs a été contrôlé avec un colorimètre Eye One Display 2 (GretagMacbeth) de manière à ce qu'il n'existe aucune variation chromatique entre les différents stimuli.

5.3.3 Stimuli et conditions de visualisation

Les tests se sont déroulés dans une salle de tests psychophysiques dont les paramètres d'éclairage sont contrôlés. Les murs de cette salle sont recouverts de tissu mat d'une couleur neutre D65. L'éclairage ambiant a été fixé à 10 cd/m².

La distance de visualisation a été fixée à 144 cm, la taille de l'écran dans le champ visuel de l'observateur était alors de 16.9 degrés visuels horizontalement et 12.7 verticalement. Dans ces conditions, la taille d'un pixel est inférieure à 1 minute d'angle visuel. L'écran de test présenté aux observateurs est découpé verticalement en deux parties égales (figure 5.1). Le stimulus stationnaire est affiché dans la partie supérieure de l'écran tandis que le stimulus en mouvement est affiché dans la partie inférieure.

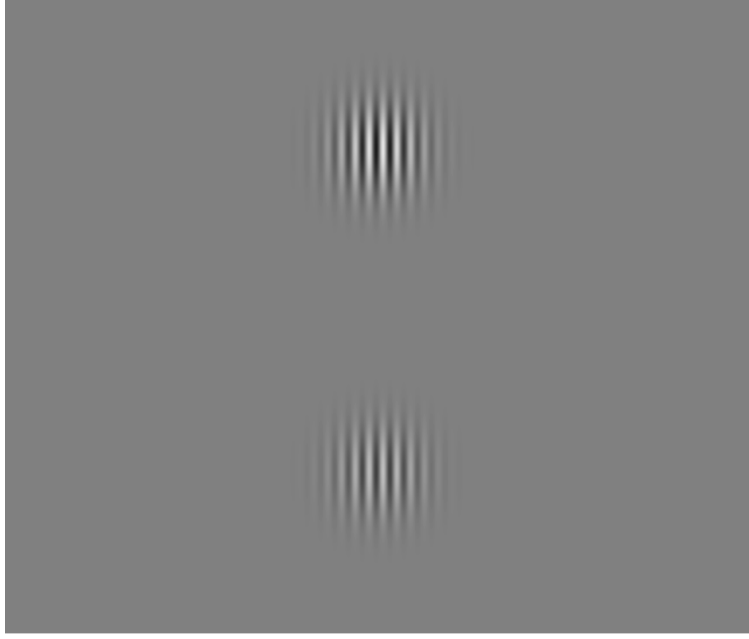


FIGURE 5.1 – Écran présenté aux observateurs durant le test : sinusoïde stationnaire en haut et sinusoïde en mouvement en bas. Les deux stimulus sont vus à travers une fenêtre gaussienne stationnaire d’une largeur de 2 degrés visuels.

Les stimuli ont été générés avec Matlab, à l’aide de l’extension PsychToolbox [Brainard 97]. Ils sont constitués d’une sinusoïde orientée horizontalement vue à travers une fenêtre gaussienne stationnaire d’une largeur à mi-hauteur de 2 degrés visuels. Le contraste d’un stimulus est défini comme la modulation de la sinusoïde, c’est à dire :

$$C = \frac{\Delta L}{L_{\text{moy}}} \quad (5.4)$$

ΔL est l’amplitude de la sinusoïde et L_{moy} la luminance moyenne de l’écran. Le contraste du stimulus stationnaire est noté C_s , il sera ajusté au fur et à mesure des réponses de l’observateur durant le test. Le contraste du stimulus en mouvement, noté C_m , reste fixe.

Le stimulus en mouvement est spatialement périodique. Par conséquent, lorsqu’il se déplace, la luminance en un point donné de l’écran varie de façon périodique avec le temps. La fréquence temporelle correspondante dépend à la fois de la vitesse du stimulus et de sa fréquence spatiale. La relation entre les trois paramètres du stimulus en mouvement est la suivante :

$$F_t = F_s \cdot V \quad (5.5)$$

Le tableau 5.1 présente les différentes valeurs de la fréquence temporelle F_t et de la fréquence spatiale F_s du stimulus en mouvement qui ont été testées durant cette expérience. Les valeurs présentées dans les cellules du tableau correspondent à la vitesse de translation V correspondante (d’après la relation 5.5). Pour tous les couples (F_t, F_s) testés durant l’expérience, la fréquence spatiale du stimulus stationnaire est toujours égale à F_s . De cette manière, le contenu spatial des deux stimuli

		Fréquence spatiale F_s (cpd)				
		0.8	1.6	3.2	6.4	12.8
Fréquence temporelle F_t (Hz)	2.4	3	1.5	-	-	-
	4.8	6	3	1.5	0.75	-
	9.6	12	6	3	1.5	0.75
	19.2	24	12	6	3	1.5

TABLE 5.1 – Combinaison des différentes fréquences spatiales F_s et fréquences temporelles F_t du stimulus en mouvement testées durant l'expérience. Les valeurs dans les cellules du tableau correspondent à la vitesse de translation V correspondante.

est le même et la seule différence intervient dans le domaine temporel.

De manière à ce que le stimulus se déplace relativement au référentiel rétinien, il a été demandé aux observateurs de fixer le centre du stimulus, sans le suivre des yeux. La fenêtre stationnaire, associée à des vitesses relativement élevées, a permis de respecter cette condition de manière aisée. Il n'a donc pas été nécessaire de rajouter un point de fixation au centre du stimulus, ce qui aurait pu interférer avec la perception du motif. Cependant, pour de faibles vitesses et de hautes fréquences spatiales, le suivi du stimulus intervient inévitablement. Pour cette raison, nous n'avons pas testé ces combinaisons. Elles sont marquées d'un tiret dans le tableau 5.1.

Enfin, lors de la présentation des stimuli, ceux-ci sont affichés à l'écran progressivement selon un demi-cycle sinusoïdal, il en va de même pour leur disparition. Cette précaution permet d'éviter un changement temporel brutal qui pourrait affecter la perception. Entre chaque présentation, un écran uniforme de luminance égale à la luminance moyenne est présenté durant 2 secondes.

5.3.4 Procédure

Pour obtenir une évaluation quantitative de l'impression subjective du contraste des stimuli, nous avons utilisée une procédure à deux alternatives et choix forcé (2AFC). Les deux stimuli sont présentés simultanément à l'observateur pendant 5 secondes. À la fin de la présentation, il doit choisir lequel des deux lui est apparu comme ayant le plus fort contraste.

Durant chaque expérience, les paramètres du stimulus en mouvement (fréquence spatiale F_s , vitesse V et contraste C_m) sont fixés. Celui-ci est présenté simultanément avec le stimulus stationnaire de même fréquence spatiale. Le contraste du stimulus stationnaire est initialement fixé à 1. Après chaque présentation, l'observateur indique en pressant un bouton quel stimulus possède le plus fort contraste, selon sa perception. Si le stimulus en mouvement est choisi, le contraste du stimulus stationnaire est augmenté d'un pas dans la présentation suivante. Au contraire, si le stimulus stationnaire est choisi, alors son contraste est diminué d'un pas dans la présentation suivante.

Le pas de variation du contraste du stimulus stationnaire est adapté durant l'expérience. Il est égal à 0.1 en début d'expérience et ce tant que le contraste est diminué (trajet descendant). Lorsque le stimulus en mouvement est choisi pour la première fois, le pas devient égal à 0.05, et reste à cette valeur tant que le contraste est augmenté (trajet ascendant). Ces deux premiers trajets permettent

de couvrir rapidement la dynamique de contraste pour converger vers la zone d'intérêt. Le pas est ensuite réduit à 0.01 pour les deux trajets suivants (descendant et ascendant), et encore une fois à 0.004 pour les deux derniers trajets (descendant et ascendant). Après six changements de direction, on considère que l'expérience est terminée. La valeur moyenne des deux derniers extrema est prise comme valeur du contraste égalisé C_s .

L'égalisation de contraste a été effectuée pour 16 stimuli en mouvement différents (tableau 5.1), à chaque fois pour quatre valeurs de contraste C_m : 0.08, 0.15, 0.30 et 0.50. Au total, ce sont donc 64 stimuli différents qui ont été évalués. Les 64 expériences d'égalisation de contraste ont été répétées deux fois par chacun des quatre observateurs. Seize sessions, comportant huit expériences chacune, ont été effectuées par chaque observateur, à raison d'une session par jour au maximum. La durée d'une session était de l'ordre de 20 minutes. La durée totale de l'expérimentation fut donc d'environ 20 heures pour chacun des observateurs (réparties sur trois semaines).

5.4 Résultats

5.4.1 Données mesurées

Pour chaque expérience d'égalisation de contraste, *i.e.* pour chaque combinaison de paramètres (F_s, F_t, C_m) , on obtient un couple de contraste (C_s, C_m) tel que le contraste du stimulus stationnaire égalisé C_s est perçu d'une magnitude équivalente au contraste du stimulus en mouvement C_m :

$$C_m(F_s, F_t) \equiv C_s(F_s) \quad (5.6)$$

La figure 5.2 présente le contraste du stimulus stationnaire C_s en fonction du contraste du stimulus en mouvement C_m , pour chaque stimulus. Les valeurs sont les moyennes des résultats obtenus par les cinq observateurs, les écarts-types sont représentés par les barres verticales. Pour la majorité des observateurs, l'égalisation n'a pas été possible pour les trois stimuli correspondant aux couples $(F_s = 0.8 \text{ cpd}, F_t = 19.2 \text{ Hz})$, $(F_s = 1.6 \text{ cpd}, F_t = 19.2 \text{ Hz})$ et $(F_s = 12.8 \text{ cpd}, F_t = 9.6 \text{ Hz})$, car ceux-ci n'étaient pas ou peu visibles.

Le contraste du stimulus stationnaire C_s , ajusté par l'observateur, est perceptuellement égal au contraste perçu du stimulus en mouvement. Par opposition, le contraste du stimulus en mouvement C_m correspond au contraste « physique » du stimulus. De manière à évaluer l'influence du mouvement sur la perception du contraste, le rapport du contraste du stimulus stationnaire C_s (contraste perçu) sur le contraste du stimulus en mouvement C_m (contraste physique) est calculé. Si ce rapport est supérieur à 1, cela signifie que le contraste perçu est plus fort que le contraste physique du stimulus en mouvement. En d'autres termes, le mouvement implique une augmentation de la sensibilité. Dans le cas contraire, si le rapport est inférieur à 1, alors le contraste physique du stimulus en mouvement était plus élevé lors de l'égalisation : le mouvement est alors responsable d'une diminution de la sensibilité.

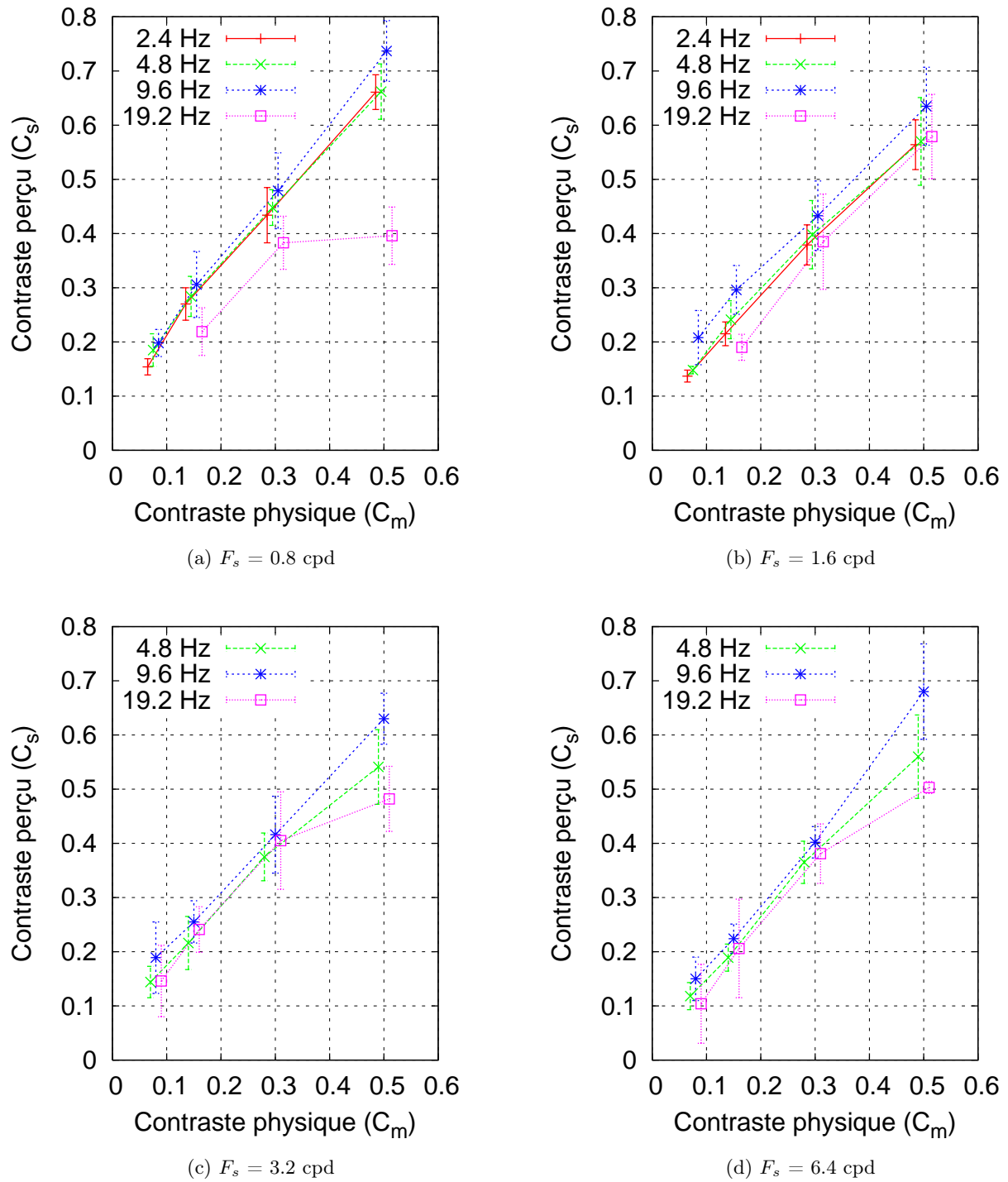


FIGURE 5.2 – Résultats de l'égalisation de contraste : contraste perçu (contraste du stimulus stationnaire C_s ajusté par l'observateur) en fonction du contraste physique C_m . La fréquence spatiale des deux stimuli vaut 0.8 cpd (a), 1.6 cpd (b), 3.2 cpd (c), et 6.4 cpd (d). La fréquence temporelle du stimulus en mouvement varie d'une courbe à l'autre (cf. légende). Les barres verticales présentent l'écart-type sur la valeur d'ajustement du contraste stationnaire. Pour chacune des quatre valeurs de contraste en mouvement C_m les points ont été décalés horizontalement pour faciliter la lecture des courbes.

Ce rapport peut également être exprimé comme le rapport de la sensibilité au stimulus en mouvement sur la sensibilité au stimulus stationnaire (la sensibilité est définie comme l'inverse du contraste). Il peut donc être vu comme le gain en sensibilité apporté par le mouvement. On le note G , il dépend à la fois de la fréquence spatiale F_s des deux stimuli et de la fréquence temporelle F_t du stimulus en mouvement :

$$G(F_s, F_t) = \frac{1/C_m(F_s, F_t)}{1/C_s(F_s)} \quad (5.7)$$

Le gain en sensibilité constitue la réponse supraliminaire du système visuel humain aux stimuli en mouvement. La figure 5.3 présente le gain (moyenne sur les quatre observateurs) en fonction du contraste du stimulus stationnaire C_s , pour chacun des 16 stimuli. On observe que pour la majorité des stimuli, et sur la plage de contraste considérée, le gain est supérieur à 1, ce qui indique que le mouvement est responsable d'un accroissement du contraste perçu. De plus, il décroît au fur et à mesure que le contraste physique du stimulus augmente. L'influence de la vitesse est donc de moins en moins forte au fur et à mesure que l'on s'éloigne du seuil de perception.

5.4.2 Modélisation du gain

De manière à pouvoir étudier la variation du gain en fonction de la fréquence spatiale F_s , de la fréquence temporelle F_t et du contraste physique C_m , nous avons effectué une modélisation prenant en compte ces quatre paramètres.

Cette modélisation a été réalisée en plusieurs étapes. Nous avons tout d'abord effectué un ajustement de courbe sur les valeurs de gain en fonction du contraste physique du stimulus (figure 5.3). Seize ajustements de courbes ont été réalisés, sur chacun des 16 couples (F_s, F_t) mesurés. Chacune de ces modélisations a utilisé la fonction suivante :

$$G_{F_s, F_t}(C_m) = a_{F_s, F_t} + b_{F_s, F_t} \cdot \exp(-c_{F_s, F_t} \cdot C_m) \quad (5.8)$$

Ce modèle présente trois paramètres, a_{F_s, F_t} , b_{F_s, F_t} et c_{F_s, F_t} , qui ont été optimisés indépendamment pour chacun des 16 couples (F_s, F_t) . Nous avons donc 16 triplets (a, b, c) .

La deuxième étape de cette modélisation a consisté à déterminer l'évolution des paramètres (a, b, c) en fonction de la fréquence temporelle F_t et de la fréquence spatiale F_s . Ici encore, nous avons réalisé un ajustement de courbe pour chacun des trois paramètres (a, b, c) . Cette modélisation empirique a permis d'obtenir les modèles suivants :

$$\begin{aligned} a(F_t) &= p_1 + p_2 \cdot F_t \\ b(F_t) &= p_3 \cdot \exp\left(\frac{-(F_t - p_4)^2}{2 \cdot p_5^2}\right) \\ c(F_s, F_t) &= p_6 \cdot \exp\left(\frac{-F_t}{p_7}\right) \cdot F_s + p_8 \cdot \exp\left(\frac{-F_t}{p_9}\right) + p_{10} \end{aligned}$$

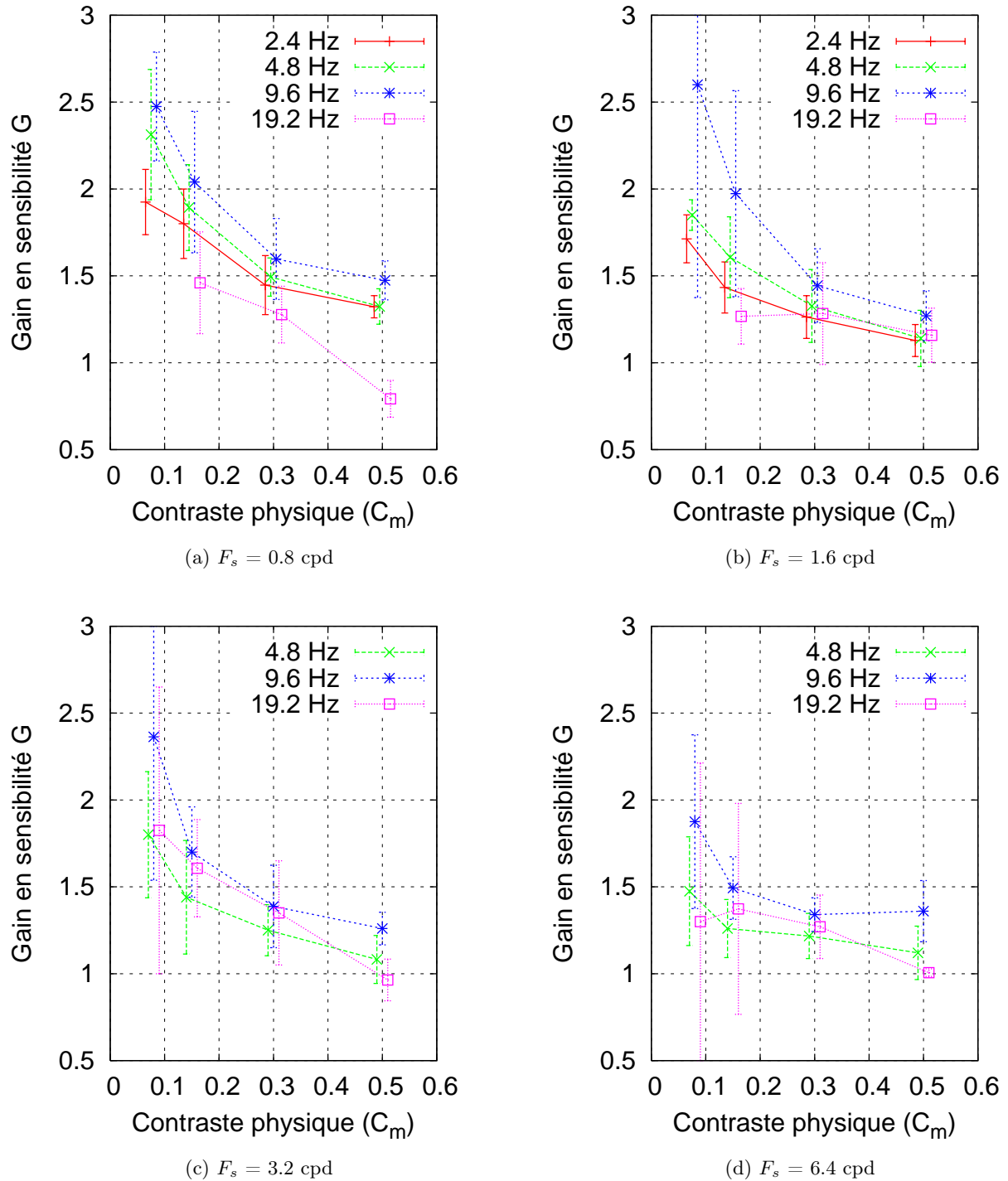


FIGURE 5.3 – Gain en sensibilité G en fonction du contraste physique du stimulus en mouvement C_m , pour des stimuli d’une fréquence spatiale égale à 0.8 cpd (a), 1.6 cpd (b), 3.2 cpd (c), et 6.4 cpd (d). La fréquence temporelle du stimulus en mouvement varie d’une courbe à l’autre (cf. légende). Les barres verticales présentent l’écart-type sur la valeur d’ajustement du contraste stationnaire. Pour chacune des quatre valeurs de contraste en mouvement C_m les points ont été décalés horizontalement pour faciliter la lecture des courbes.

p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}
1.127	-0.01652	2.232	11.49	6.420	4.186	5.320	-6.892	3.400	3.434

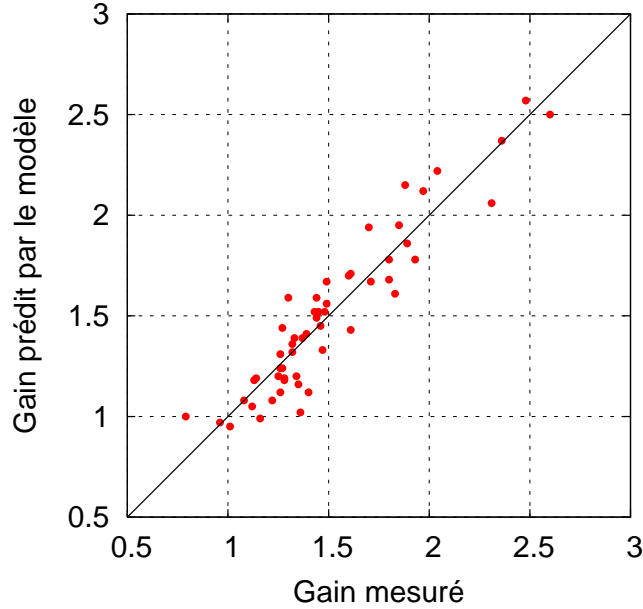
TABLE 5.2 – Valeurs des constantes nécessaires à la modélisation des paramètres a , b et c .

FIGURE 5.4 – Comparaison entre les valeurs de gain obtenues par l'expérience et celles prédites par le modèle décrit par la relation 5.9.

Les valeurs des constantes p_1 à p_{10} sont données dans le tableau 5.2.

Le modèle global de prédiction du gain de sensibilité est donc le suivant :

$$G(F_s, F_t, C_m) = p_1 + p_2 \cdot F_t + p_3 \cdot \exp\left(\frac{-(F_t - p_4)^2}{2 \cdot p_5^2}\right) \cdot \exp\left(-\left(p_6 \cdot \exp\left(\frac{-F_t}{p_7}\right) \cdot F_s + p_8 \cdot \exp\left(\frac{-F_t}{p_9}\right) + p_{10}\right) \cdot C_m\right) \quad (5.9)$$

La figure 5.4 compare les valeurs de gain prédites par le modèle avec celles obtenues à partir des données mesurées lors des expériences d'égalisation de contraste. Le coefficient de corrélation linéaire entre les deux séries de points est de 0.9383 et l'écart quadratique moyen est de 0.020.

5.4.3 Surface spatio-temporelle du gain en sensibilité

La surface spatio-temporelle du gain en sensibilité est tracée figure 5.5 dans le domaine fréquentiel, pour un contraste physique de référence égal à 0.1. La réponse supraliminaire est de forme passe-

bande le long de l'axe des fréquences temporelles avec un pic de sensibilité qui se situe entre 10 Hz (pour les faibles fréquences spatiales) et 15 Hz (pour les hautes fréquences spatiales). Le long de l'axe des fréquences spatiales, la réponse supraliminaire est de forme passe-bas, quelque soit la fréquence temporelle. La figure 5.6 présente plusieurs coupes de la surface spatio-temporelle du gain, pour des valeurs données de la fréquences temporelle F_t (5 Hz et 15 Hz) et de la valeur spatiale F_s (5 cpd et 15 cpd). Les courbes sont tracées pour plusieurs valeurs du contraste physique, en plus de la valeur 0.1 déjà présentée sur la figure 5.5. On remarque que la forme de la réponse supraliminaire s'aplatit au fur et à mesure que le contraste de référence augmente, aussi bien le long de l'axe des fréquences spatiales que le long de l'axe des fréquences temporelles. Ce comportement est similaire à celui obtenu pour des stimuli fixes [Georgeson 75] et pour des stimuli clignotants [Bowker 83] : l'influence de la fréquence spatiale et de la fréquence de clignotement est moins importante à mesure que l'on s'éloigne du seuil. De plus, le gain en sensibilité décroît à mesure que le contraste physique augmente, indiquant un effet de « stabilisation du contraste ».

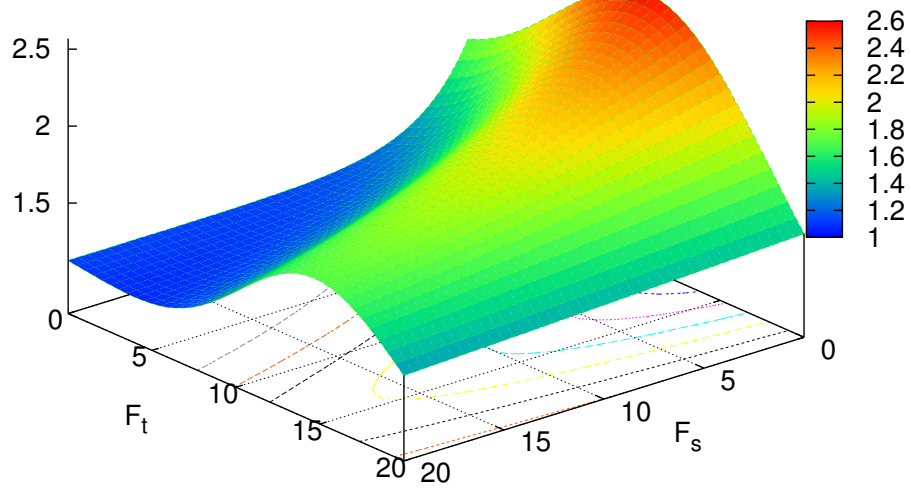
5.5 Discussion

5.5.1 Stabilisation du contraste

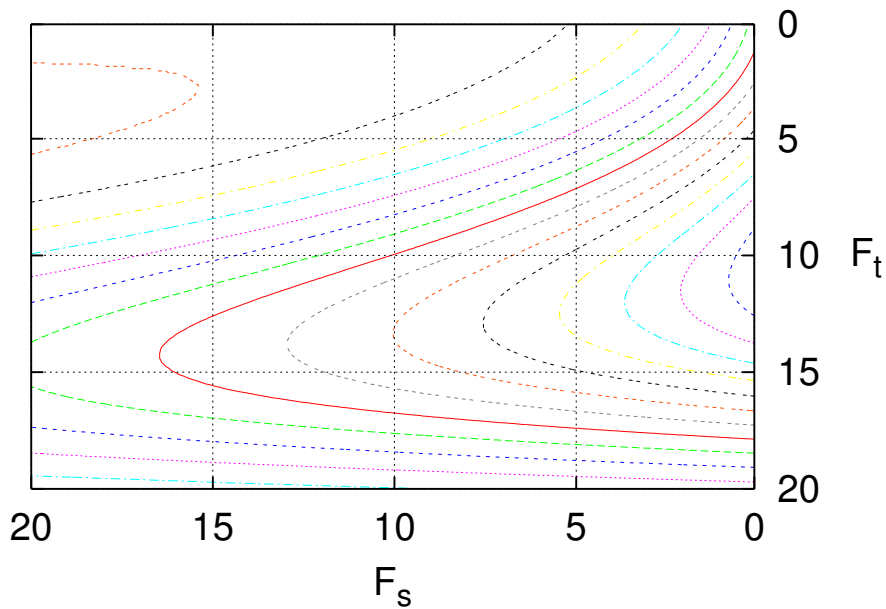
Le premier résultat de cette expérience est que la réponse supraliminaire à un stimulus en mouvement devient de plus en plus uniforme à mesure que le contraste physique du stimulus augmente. Ce résultat est en accord avec les travaux précédents, lesquels décrivent le même phénomène tant en fréquences temporelles [Georgeson 87, Waugh 94, Manahilov 03, Fiser 03, Bex 07] qu'en fréquences spatiales [Bryngdahl 66, Watanabe 68, Georgeson 75, Cannon 85]. Plus particulièrement, la non-linéarité de la réponse du système visuel humain est illustrée par le fait que le contraste perçu pour des stimuli de hautes et basses fréquences spatiales augmente plus rapidement que celui des stimuli de fréquence spatiale proche du pic des CSF. Ce phénomène est appelé la « stabilisation du contraste » (*contrast constancy*).

Pour expliquer cet effet, Georgeson & Sullivan [Georgeson 75] ont supposé qu'il existait dans le cortex visuel un processus qui compense l'atténuation des hautes et basses fréquences spatiales en ajustant un gain permettant d'amplifier ce type de signaux. Une réduction de flou est alors effectuée sur le signal pour optimiser la vision. Il semblerait d'après les résultats de nos mesures que ce phénomène existe également pour les stimuli en mouvement. C'est un résultat intéressant dans l'optique de l'étude de la perception du flou de mouvement sur LCD. En effet, comme on l'a vu précédemment un objet en mouvement sur un écran LCD se déplace sur la rétine. Les hautes fréquences spatiales du signal sont alors atténuées (flou de bougé). Cependant l'existence d'un tel mécanisme pourrait réduire cette atténuation des hautes fréquences, ce qui signifie, si cette hypothèse est vérifiée, que *le flou de mouvement perçu serait moins important que le flou de mouvement mesuré*.

Gain en sensibilité



(a) Surface du gain en sensibilité.



(b) Topographie de la surface, les valeurs des courbes de niveau vont de 1.1 à 2.5 avec un incrément de 0.1.

FIGURE 5.5 – Surface et topographie du gain en sensibilité G dans le domaine des fréquences spatio-temporelles, tracés à partir du modèle décrit par la relation 5.9, pour un contraste de 0.1.

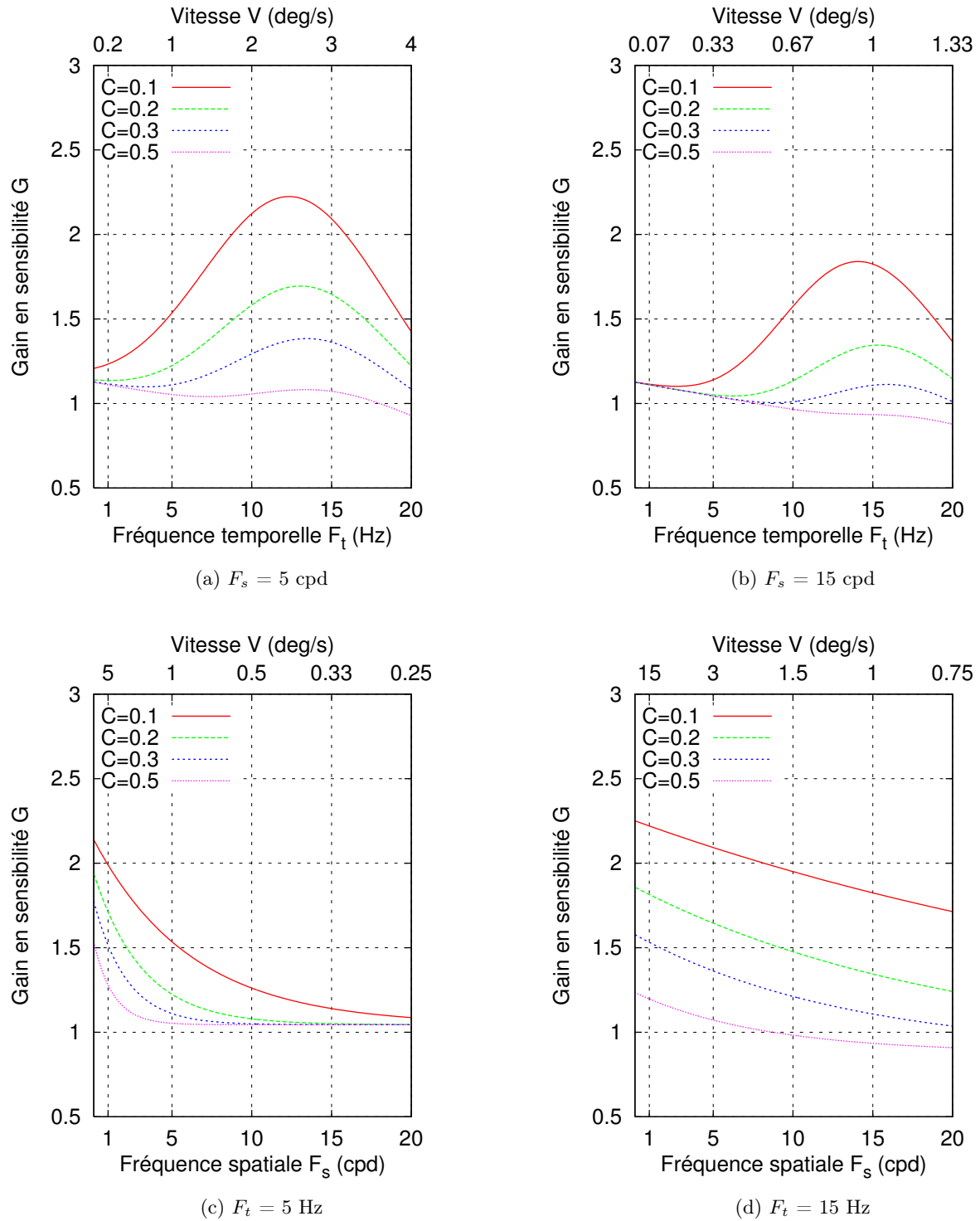


FIGURE 5.6 – Gain en sensibilité G en fonction de la fréquence temporelle F_t (la vitesse V est également indiquée sur l'axe horizontal supérieur) pour une fréquence spatiale F_s de 5 cpd (a), 15 cpd (b). Également en fonction de la fréquence spatiale F_s pour une fréquence temporelle F_t fixée à 5 Hz (c), et 15 Hz (d). Chaque courbe correspond à une valeur différente du contraste physique du stimulus (cf. légende).

5.5.2 Comparaison avec les mesures effectuées au seuil de visibilité

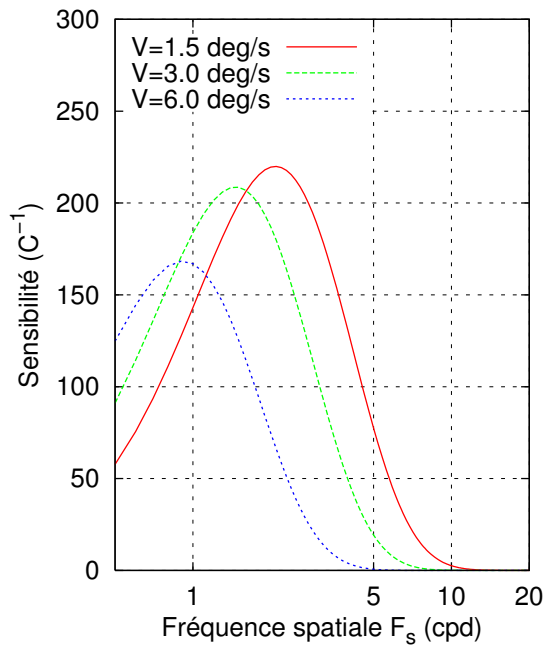
Les mesures du seuil différentiel de visibilité pour des stimuli en mouvement (par exemple celles de Koenderink [Koenderink 79] et de Kelly [Kelly 84]) présentent une forme passe-bande le long de l'axe des fréquences temporelles pour les faibles fréquences spatiales (inférieures à 3 cpd) avec un maximum de sensibilité aux alentours de 6 Hz. La surface de détection est également de forme passe-bande le long de l'axe des fréquences spatiales pour de faibles fréquences temporelles (inférieures à 6 Hz) avec un maximum de sensibilité aux alentours de 3 cpd. Au delà, la surface est de forme passe-bas, aussi bien le long des fréquences spatiales que le long des fréquences temporelles.

La réponse supraliminaire du SVH au stimuli en mouvement que nous venons de mesurer est sensiblement différente des résultats obtenus par Koenderink et Kelly. On retrouve la forme passe-bas le long de l'axe des fréquences temporelles, mais avec un maximum entre 10 et 15 Hz. Ce maximum se décale légèrement vers les hautes fréquences temporelles lorsque la fréquence spatiale du stimulus augmente. Contrairement à ce qui se passe au seuil de détection, la forme passe-bas le long de l'axe des fréquences temporelles est toujours observée pour des hautes fréquences spatiales (cf. figure 5.6b). Le long de l'axe des fréquences spatiales, la forme de la réponse est passe-bas quelle que soit la fréquence temporelle (cf. figure 5.6c, d).

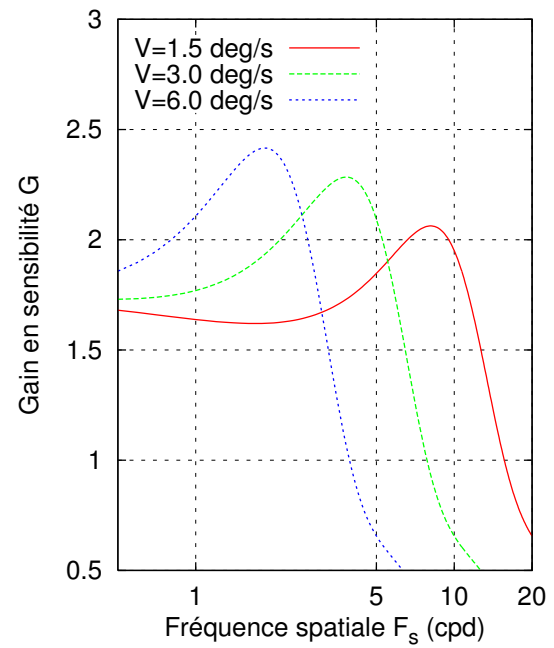
Les fonctions de sensibilité au contraste en mouvement déterminées par Kelly [Kelly 79b] sont représentées figure 5.7a pour trois valeurs fixes de vitesse. Elles montrent que pour une vitesse donnée, un intervalle de l'espace des fréquences spatiales est privilégié par le système visuel. Cet intervalle se déplace vers les basses fréquences spatiales à mesure que la vitesse augmente, maintenant le maximum de sensibilité à une fréquence temporelle d'environ 6 Hz. De plus, le maximum de ces fonctions de sensibilité décroît avec la vitesse. En d'autres termes, plus la vitesse est élevée, plus les basses fréquences spatiales sont privilégiées par le SVH, mais avec une sensibilité moindre.

La modélisation du gain de sensibilité aux stimuli en mouvement nous permet de représenter le même type d'information à des niveaux supraliminaires et pour une vitesse constante. Nous obtenons alors les courbes de la figure 5.7 (b, c et d) qui correspondent chacune à une valeur donnée du contraste physique du stimulus en mouvement. Comme pour les données recueillies au seuil différentiel de visibilité par Kelly, un effet d'extraction des fréquences spatiales peut être observé. Chaque réponse supraliminaire à vitesse constante présente un maximum qui est déplacé vers les basses fréquences spatiales à mesure que la vitesse du mouvement augmente. Il est intéressant de noter la similarité entre les résultats obtenus au seuil de visibilité et ceux obtenus à des niveaux supraliminaires concernant la sélection d'une bande donnée de fréquences spatiales.

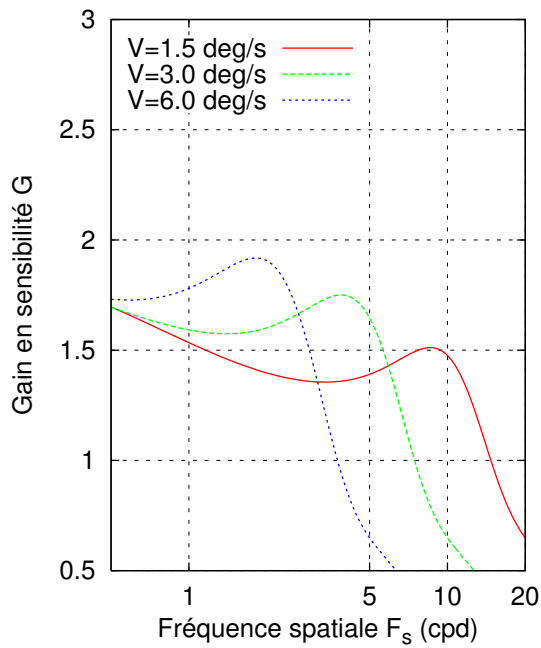
Cependant, pour chaque bande de fréquences sélectionnée et contrairement à ce qui se passe au seuil de visibilité, le gain en sensibilité dû au mouvement est plus important pour des vitesses plus élevées. Ce résultat peut laisser supposer la présence d'un processus de compensation lors de la perception de stimuli en mouvement, pour obtenir une stabilisation du contraste à des niveaux supraliminaires. En effet, la sensibilité au seuil diminue à mesure que la vitesse augmente, tandis que le gain en sensibilité à des niveaux supraliminaires augmente avec la vitesse. Une adaptation à la vitesse semble donc s'exercer au delà du seuil.



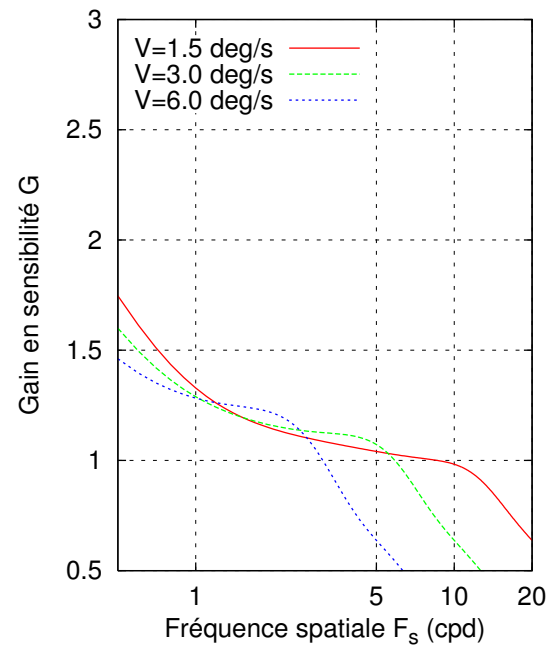
(a) Réponse au seuil différentiel de visibilité



(b) Réponse supraliminaire pour un contraste physique de 0.1



(c) Réponse supraliminaire pour un contraste physique de 0.2



(d) Réponse supraliminaire pour un contraste physique de 0.5

FIGURE 5.7 – Fonction de sensibilité au contraste en mouvement mesurée au seuil différentiel de visibilité (tirée des travaux de Kelly [Kelly 79b]) (a) et réponses supraliminaires en fonction de la fréquence spatiale, pour des contrastes de référence de (a) 0.1, (b) 0.2, et (c) 0.5, et pour trois vitesses (cf. légende).

5.6 Conclusion

Dans ce chapitre, des tests psychophysiques ont été menés de manière à mesurer la réponse du système visuel humain à des stimuli en mouvement sur la rétine. Ces expériences ont été menées à des niveaux de contraste supraliminaires. La procédure choisie consistait en une égalisation de contraste entre un stimulus fixe et un stimulus en mouvement de même fréquence spatiale. Le contraste du stimulus fixe était variable, et un protocole de décision par choix forcé faisait tendre cette valeur de contraste vers l'isoperception avec le contraste du stimulus en mouvement. Au final, plus de 60 couples de stimuli ont été testés. Pour chacun d'entre eux, le gain en sensibilité dû au mouvement a été calculé comme le rapport entre le contraste du stimulus fixe (égal au contraste perçu du stimulus en mouvement) et le contraste du stimulus en mouvement (contraste physique). Sur la majorité du domaine spatio-temporel exploré, ce gain est supérieur à 1, indiquant que le mouvement est responsable d'une augmentation du contraste perçu.

Ce gain en sensibilité a été modélisé par l'intermédiaire d'une fonction à trois variables (fréquence temporelle, fréquence spatiale et contraste physique). Cette modélisation a permis d'étudier la surface spatio-temporelle de cette réponse supraliminaire du SVH à des stimuli en mouvement. En particulier, nous avons montré que le gain en sensibilité diminue, et qu'il est moins dépendant des caractéristiques spatio-temporelles du stimulus, à mesure que le contraste physique de ce dernier augmente. Ce résultat tend à prouver que le phénomène de « stabilisation du contraste » déjà mise en évidence sur des stimuli supraliminaires fixes ou clignotants, intervient également pour les stimuli en mouvement. La chute de sensibilité aux hautes fréquences spatiales et temporelles observée au seuil différentiel de visibilité est donc atténuée pour des valeurs de contraste supraliminaires. Ce rehaussement des hautes fréquences devrait se traduire dans le cadre de notre étude par une réduction du flou perçu sur les contours en mouvement.

Enfin, des similarités ont été trouvées entre la réponse supraliminaire du SVH et les fonctions de sensibilité aux stimuli en mouvement en seuil de détection. Notamment, pour une vitesse de déplacement donnée certaines fréquences spatiales sont privilégiées par le système visuel (filtre passe-bande étroit). Néanmoins, les paramètres de ces filtres sont différents entre le seuil et les niveaux supraliminaires. Ici encore, il semblerait qu'une étape de compensation non-linéaire intervienne puisque la sensibilité augmente avec la vitesse à des niveaux de contraste supraliminaires, tandis qu'au seuil de visibilité, elle diminue lorsque la vitesse augmente. Ce résultat nous permet de supposer que la perception du flou de mouvement ne dépend pas de la vitesse de déplacement du contour. La réduction du flou de mouvement supposée précédemment devrait donc s'appliquer dans les mêmes proportions, quelle que soit la vitesse de déplacement.

Les résultats présentés dans ce chapitre constituent une contribution au domaine de la vision appliquée. Ces expériences ont principalement été motivées par le fait que nous n'avons trouvé aucune étude dans la littérature concernant la caractérisation de la réponse du système visuel à des stimuli en mouvement et pour des niveaux de contraste supraliminaires. Les résultats obtenus permettent d'éclairer le comportement « haut-niveau » du système visuel dans certaines conditions particulières. Cependant, dans le cadre de notre étude sur la perception du flou de mouvement, l'utilisation de ces résultats s'est révélée délicate. En effet, le flou de mouvement apparaît sur des contenus en

mouvement, généralement des contours, qui se prêtent mal à l'analyse fréquentielle. Néanmoins, la modélisation de la réponse supraliminaire du SVH à des stimuli en mouvement nous a tout de même permis d'énoncer une hypothèse forte : le flou de mouvement perçu par un observateur est moins important que le flou de mouvement mesuré (formé sur la surface rétinienne).

Pour vérifier cette hypothèse, nous avons décidé de mener de nouvelles expérimentations, en mesurant cette fois-ci directement le flou de mouvement perçu par les observateurs sur des stimuli simples. Cette seconde étude psychovisuelle est le sujet du chapitre suivant.

Chapitre 6

Perception du flou de mouvement

Sommaire

6.1	Introduction	103
6.2	Travaux antérieurs	104
6.3	Description de l'expérience	105
6.3.1	Mesures physiques	105
6.3.2	Protocole	105
6.3.3	Écrans testés	107
6.3.4	Conditions de visualisation	107
6.3.5	Observateurs	108
6.4	Résultats et interprétation	108
6.4.1	Résultats	108
6.4.2	Comparaison avec les autres travaux	109
6.4.3	Étude de la variation du coefficient A	111
6.4.4	Évaluation globale du modèle	115
6.5	Vers une mesure perceptive du flou de mouvement	117
6.6	Conclusion	118

6.1 Introduction

Les expérimentations menées dans le chapitre précédent ont permis d'énoncer l'hypothèse selon laquelle le flou de mouvement perçu serait moins important que le flou de mouvement mesuré de manière optique. Cependant, ce résultat qualitatif est difficilement utilisable tel quel pour caractériser quantitativement le flou de mouvement perçu par les observateurs. Il est donc nécessaire d'explorer la perception du flou de mouvement de façon plus directe pour mesurer la valeur du flou perçu sur un contour en mouvement.

De manière à répondre à ce problème, nous avons développé un protocole de mesure du flou perçu sur un contour en mouvement. Ce protocole s'inspire des travaux antérieurs présents dans la littérature [Feng 06, Teunissen 06, Someya 07]. Il consiste à comparer le flou de mouvement perçu sur un

contour en déplacement avec un flou simulé à partir des résultats obtenus dans le chapitre 4. Le flou simulé est donc obtenu à partir de la réponse indicielle des pixels de l'écran testé. La principale différence avec les travaux existants réside dans le nombre de conditions testées. En effet, trois écrans ont été utilisés, 20 transitions gris-à-gris ont été testées sur chaque écran, et entre quatre et neuf vitesses différentes ont été testées pour chaque transitions. En comparaison, les trois études antérieures ne concernaient qu'un seul écran et seulement quelques transitions et/ou vitesses, ce qui nous a décidé à mener nos propres expériences en explorant un ensemble de conditions plus large.

Dans ce chapitre, le protocole de notre expérience est décrit et les résultats obtenus sont présentés. Ces derniers sont ensuite commentés et comparés aux résultats existants. Enfin, nous développons un modèle perceptif permettant de calculer le flou de mouvement perçu à partir des valeurs de flou mesuré objectivement. Ce modèle nous permet de définir une mesure perceptive globale du flou de mouvement perçu sur un écran donné.

6.2 Travaux antérieurs

Trois études ont été publiées ces dernières années concernant la mesure subjective du flou de mouvement perçu. Someya et Sugiura [Someya 07] ont développé des tests d'égalisation de flou de mouvement entre un contour en mouvement sur LCD et un contour en mouvement sur CRT sur lequel le flou était simulé. Dans cette étude, 42 transitions ont été testées, mais avec une seule valeur de vitesse de déplacement du contour. Les résultats ont montré une bonne corrélation entre le flou mesuré et le flou perçu. Néanmoins, la complexité du dispositif expérimental et plus particulièrement les différences physiques entre CRT et LCD ne permettent pas de tirer des conclusions définitives. Teunissen *et al.* [Teunissen 06] ont également développé une expérience d'égalisation du flou de mouvement. Les expérimentations consistaient à ajuster la vitesse d'un contour en mouvement jusqu'à ce que le flou perçu soit comparable à celui d'un contour fixe simulé à partir des mesures physiques effectuées. 8 transitions et 4 vitesses ont été testées, sur un seul écran. Une très bonne correspondance a été mesurée entre les vitesses ajustées par les observateurs et les vitesses utilisées pour simuler le flou sur le contour fixe. Enfin, Feng [Feng 06] a réalisé le même genre d'expérience avec une seule transition et 4 vitesses. La tâche des observateurs consistaient à ajuster la quantité de flou sur un contour fixe, jusqu'à ce que celle-ci corresponde à leur perception du flou sur un contour en mouvement. Les résultats obtenus sont ici légèrement différents des deux précédentes expériences. En effet, bien que la corrélation entre le flou perçu et le flou mesuré soit très bonne, l'amplitude du flou perçu est inférieure à celle du flou mesuré à l'aide d'instruments optiques.

Toutes ces études ne concernent que quelques transitions ou quelques vitesses et un seul écran chacune. De plus, des contradictions existent entre les résultats obtenus par Teunissen et Feng, sans doute en partie à cause de la différence de procédure (ajustement de la vitesse dans le premier cas, ajustement du flou dans le second cas). Nous avons donc décidé de réaliser nos propres expérimentations sur un ensemble de conditions plus large.

6.3 Description de l'expérience

6.3.1 Mesures physiques

Le flou de mouvement a été mesuré sur chaque écran selon la méthode décrite dans le chapitre 4. Le profil d'un contour en mouvement est obtenu par la convolution de la réponse indicielle temporelle de l'écran avec une fenêtre de largeur égale à la période de rafraîchissement de l'écran T . Le profil du contour est alors obtenu dans le domaine temporel et le temps de flou est mesuré entre 10% et 90% de la dynamique du contour. La valeur du temps de flou est indépendante de la vitesse de déplacement du contour.

Dans la suite de ce mémoire, on note BET_m le temps de flou obtenu à partir des mesures physiques, et BEW_m la largeur de flou calculée par la relation suivante :

$$BEW_m = BET_m \cdot V \quad (6.1)$$

Cette valeur dépend de l'écran, de la transition gris-à-gris considérée, et de la vitesse de déplacement du contour.

6.3.2 Protocole

Le protocole expérimental utilisé dans ces expérimentations est très similaire à celui décrit par Feng [Feng 06]. La tâche de l'observateur consiste à ajuster le flou d'un contour fixe jusqu'à ce qu'il corresponde à la perception du flou de mouvement sur un contour en mouvement. L'écran de test est séparé en deux parties égales comme illustré par la figure 6.1. Le contour fixe est affiché dans la partie supérieure de l'écran tandis qu'un contour en mouvement est affiché dans la partie inférieure. Les observateurs doivent suivre des yeux le contour en mouvement de manière à percevoir le flou de mouvement, puis ajuster la largeur du flou introduit sur le contour fixe. Le défilement du contour en mouvement est répété autant de fois que nécessaire, jusqu'à ce que l'observateur valide son choix. Une pause de 2 secondes est observée entre deux défilements successifs.

Pour percevoir le flou de mouvement sur le contour en déplacement, il est obligatoire de suivre des yeux ce contour. La position des deux contours est telle que lorsque que l'observateur regarde le contour fixe pour ajuster sa réponse, le contour en mouvement apparaît dans son champ de vision périphérique, sans être suivi et donc sans que le flou de mouvement ne soit perçu. De manière à ce que l'égalisation de flou se déroule correctement, nous avons bien expliqué aux observateurs que celle-ci devait se faire en deux étapes. La première étape consiste à suivre des yeux le contour en mouvement pour évaluer le flou, la seconde étape consiste à ajuster en fonction le flou sur le contour fixe. Ces deux étapes peuvent être répétées autant de fois que nécessaire mais doivent obligatoirement être effectuée séparément. Dans la pratique, l'exercice s'est révélé très aisé pour tous les observateurs.

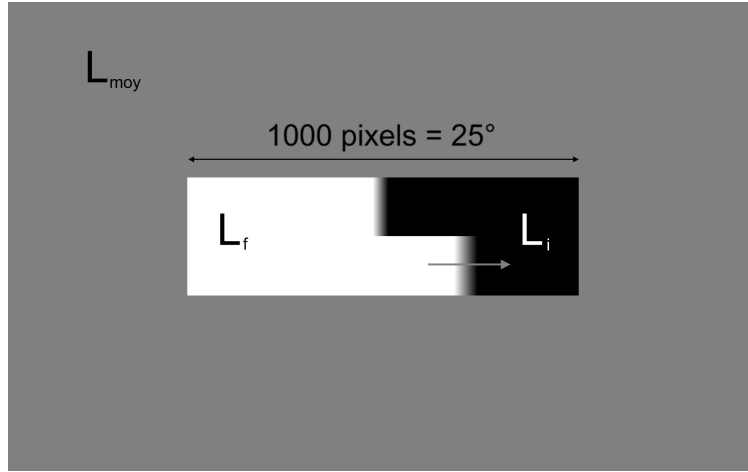


FIGURE 6.1 – Écran présenté aux observateurs durant le test. Le contour situé dans la partie inférieure est animé d'un mouvement de translation de la gauche vers la droite de vitesse V . Le contour situé dans la partie supérieure est fixe et les observateurs font varier la largeur du flou simulé sur ce contour jusqu'à ce qu'elle corresponde au flou perçu sur le contour en mouvement.

Le contour en mouvement est défini par un niveau de gris initial N_i (à la droite du contour) et un niveau de gris final N_f (à la gauche du contour). Comme cela a été décrit précédemment (cf. chapitre 4), le profil spatial du contour en mouvement peut être obtenu à partir de la réponse indicielle temporelle à la transition correspondante $N_i \rightarrow N_f$. Le contour flou fixe affiché dans la partie supérieure de l'écran est simulé à partir des résultats des mesures physiques effectuées sur l'écran testé, et pour la transition considérée.

L'observateur doit faire varier la quantité de flou sur le contour fixe via une méthode d'ajustement. Il peut donc augmenter ou diminuer la largeur du flou autant de fois que nécessaire, jusqu'à qu'il perçoive une équivalence entre la largeur de flou sur le contour fixe et sur celui en mouvement. L'ajustement est réalisé par une mise à l'échelle du profil spatial du contour obtenu à partir des mesures physiques. Le profil spatial du contour fixe peut ainsi être choisi parmi 20 profils préalablement calculés (figure 6.2).

Les différents stimuli (contour fixe et contour en mouvement) ont été générés avec Matlab, à l'aide de l'extension PsychToolbox [Brainard 97]. 20 transitions gris-à-gris ont été testées durant cette expérience :

$$N_i \rightarrow N_f \quad \text{tel que } (N_i, N_f) \in \{0, 63, 127, 191, 255\}^2 \text{ et } N_i \neq N_f \quad (6.2)$$

Chaque transition a été testée pour plusieurs vitesses de déplacement du contour. Chaque égalisation de flou a été effectuée deux fois par tous les observateurs.

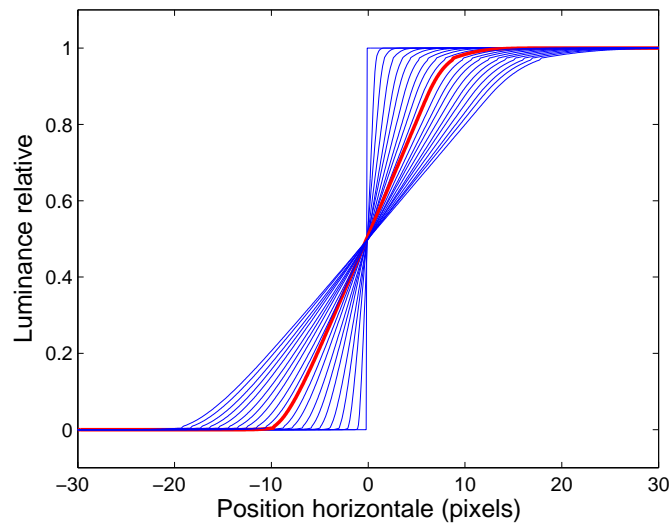


FIGURE 6.2 – Profils spatiaux des contours fixes qui peuvent être choisis par les observateurs durant la procédure d'égalisation de flou. Ils sont simulés par une mise à l'échelle du profil de flou mesuré (ligne épaisse).

Id	Type	Taille	Résolution et fréquence	Luminance maximale
1	moniteur informatique	23"	1920 × 1200 @ 60Hz	300 cd/m ²
2	moniteur informatique	30"	2560 × 1600 @ 60Hz	160 cd/m ²
3	téléviseur	37"	1920 × 1080 @ 60Hz	440 cd/m ²

TABLE 6.1 – Caractéristiques des trois écrans testés.

6.3.3 Écrans testés

L'expérience a été répétée sur trois écrans à cristaux liquides. Les caractéristiques principales de ces écrans sont mentionnées dans le tableau 6.1. La luminosité et le contraste des écrans ont été réglés conformément aux spécifications conseillées par les fabricants, un calibrage de couleur a également été effectué. L'écran 2 est équipé d'un système de réduction du flou de mouvement par clignotement du rétro-éclairage (*backlight flashing*).

La mesure du flou de mouvement selon la méthode analytique décrite dans le chapitre 4 a été réalisée sur les trois écrans testés. Et les résultats de ces mesures ont été utilisés indépendamment sur chaque écran pour simuler le flou ajustable sur le contour fixe.

6.3.4 Conditions de visualisation

Deux laboratoires ont été impliqués dans ces expérimentations : l'équipe IVC du laboratoire IRCCyN et le Video and Display Quality Lab du centre de recherche Acreo AB en Suède. Les expérimentations se sont donc déroulées dans deux salles de tests psychophysiques différentes. Les conditions de visualisation ont été précisément spécifiées de manière à obtenir des conditions expérimentales les

plus similaires possibles entre les deux salles de tests. Les paramètres d'éclairage sont contrôlés dans les deux salles et l'éclairage ambiant a été fixé à 10 cd/m^2 . Les murs des deux salles sont recouverts de tissu mat d'une couleur neutre D65.

La luminance moyenne L_{moy} de chaque écran a été choisie comme la moyenne géométrique entre la luminance minimale et la luminance maximale de l'écran : $L_{moy} = \sqrt{L_{min} \cdot L_{max}}$. Pour chaque écran, la distance de visualisation a été fixée de telle façon qu'un pixel sous-tend un angle visuel de 1.5 minutes d'arc. Cela correspond à une distance de $1.87H$ pour l'écran 1, $1.37H$ pour l'écran 2, et $2.2H$ pour l'écran 3. La zone de l'écran dans laquelle les stimuli ont été affichés était d'une taille de 1200 pixels (en horizontal) par 300 pixels (en vertical), soit 30 degrés par 7.5 degrés.

Ainsi, les conditions d'observations sont restées les mêmes pour chaque écran testé, et dans chacune des deux salles de tests.

6.3.5 Observateurs

Les écrans 1 et 3 ont été utilisés à l'IRCCyN-IVC, avec trois observateurs. L'écran 2 a été utilisé à Acreo avec trois observateurs également, différents des trois précédents.

La vision de tous les observateurs a été testée de manière à vérifier qu'ils possédaient une acuité visuelle supérieure ou égale à 9/10. Tous étaient familiers avec la procédure mise en œuvre dans les expériences après plusieurs sessions d'entraînement dont les résultats n'ont pas été pris en compte.

6.4 Résultats et interprétation

6.4.1 Résultats

La figure 6.3 illustre un exemple de résultats obtenus pour les trois écrans testés, pour les transitions $0 \rightarrow 63$ et $63 \rightarrow 191$. La droite pointillée représente le flou mesuré $BEW_m = BET_m \cdot V$ (cf. relation 6.1). Les points représentent la moyenne des valeurs de flou ajustées par les observateurs, pour chaque vitesse. Les observateurs ont effectué chaque égalisation de flou deux fois, chaque point correspond donc à la moyenne de six valeurs. L'écart-type des six valeurs est figuré par les barres verticales associées à chaque point. Nous identifions cette valeur moyenne du flou ajusté par les observateurs comme étant le flou de mouvement perçu, noté BEW_p .

Pour l'écran 1 (figures 6.3a et 6.3b), 9 vitesses ont été testées : 4, 6, 8, 10, 12, 14, 16, 18 et 20 pixels par image. À la distance de visualisation choisie, cela correspond à des vitesses respectives de 6, 9, 12, 15, 18, 21, 24, 27 et 30 degrés par secondes. Cette première série de tests a mis en évidence la très bonne stabilité des résultats de l'égalisation de flou. Sur les 20 transitions et 9 vitesses testées sur l'écran 1, la valeur moyenne de l'écart-type est de 0.8 pixels, soit 1.2 minutes d'arc, malgré seulement six occurrences sur chacune des 180 conditions de tests. Étant donné cette convergence des résultats, et devant le temps qu'il a fallu pour réaliser plus de 1000 égalisations de flou sur l'écran 1, nous avons décidé de réduire le nombre de vitesses testées sur les écrans 2 et 3. Par conséquent, seules les vitesses de 10, 15, 20 et 25 pixels par image (15, 22.5, 30 et 37.5 degrés

par seconde) ont été testées sur les écrans 2 et 3 (figures 6.3c, 6.3d, 6.3e et 6.3f). Ce qui a permis de réduire de moitié la durée totale des tests sur chaque écran.

On peut observer que, pour chaque transition et pour chaque écran, le flou de mouvement perçu varie linéairement avec la vitesse de déplacement du contour :

$$BEW_p = BET_p \cdot V \quad (6.3)$$

Par analogie avec ce qui a été montré pour le flou mesuré, le facteur de proportionnalité est de dimension temporelle. Il est identifié comme le temps de flou perçu, et noté BET_p . Cette relation est représentée par une droite sur la figure 6.3. Le coefficient de corrélation linéaire entre cette relation et les résultats subjectifs est toujours supérieur à 0.982 pour les 20 transitions testées, et sur les 3 écrans.

On peut déduire des équations 6.1 et 6.3 qu'il existe une relation de proportionnalité entre le flou de mouvement perçu et le flou de mouvement mesuré :

$$BEW_p = A \cdot BEW_m \quad (6.4)$$

avec A le coefficient de proportionnalité entre le flou de mouvement perçu et le flou de mouvement mesuré, ou encore entre le temps de flou perçu et le temps de flou mesuré :

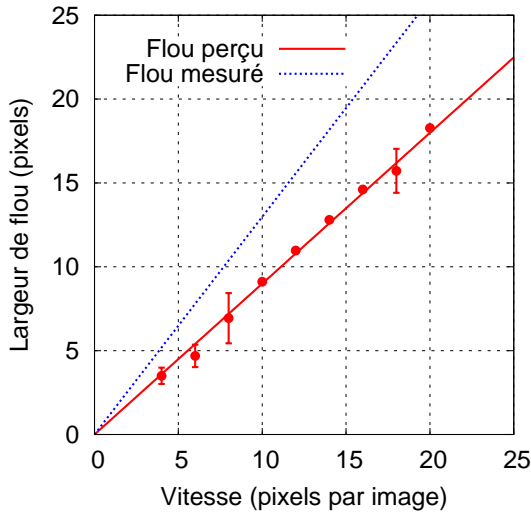
$$A = \frac{BEW_p}{BEW_m} = \frac{BET_p}{BET_m} \quad (6.5)$$

Pour la majorité des transitions testées, sur les trois écrans, ce coefficient A est inférieur à 1. Cela indique donc que le contour en mouvement est en réalité perçu de façon plus nette que celui prédit par les mesures physiques.

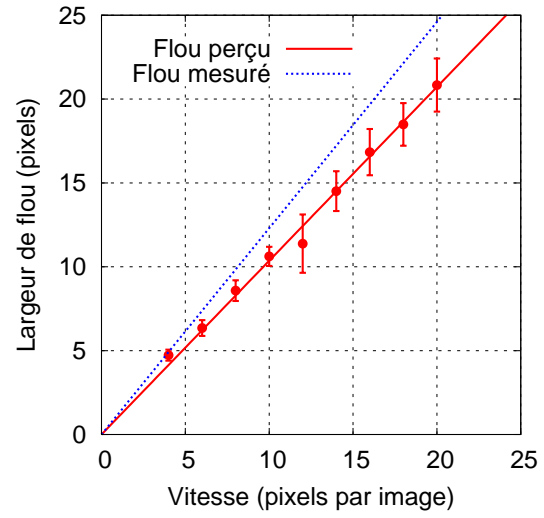
6.4.2 Comparaison avec les autres travaux

Le résultat principal de ces expérimentations est que le coefficient de proportionnalité A entre le flou perçu et le flou mesuré est presque toujours inférieur à 1 (de 0.68 à 1.03 sur toutes les transitions et pour les trois écrans, avec une valeur moyenne égale à 0.91). Ce résultat est en accord avec celui donné par Feng [Feng 06], bien que ce dernier n'ait exploré qu'une seule transition. Nous avons donc montré ici que ce phénomène apparaissait sur l'ensemble des transitions, et qu'il se répétait d'un écran à l'autre.

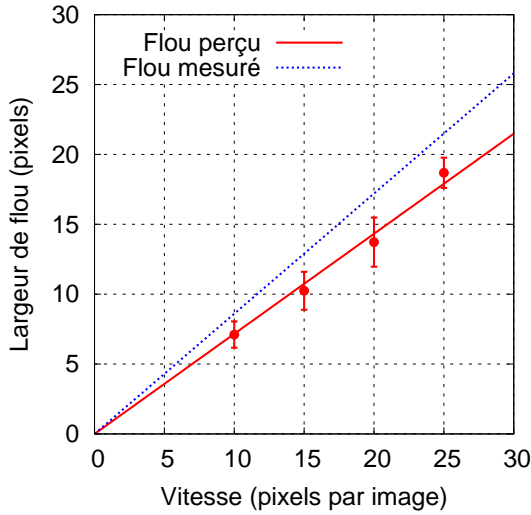
Teunissen *et al.* ont obtenu des résultats assez différents [Teunissen 06]. Les expérimentations qu'ils ont menées ont permis de tester 8 transitions avec 4 vitesses sur un seul écran. La procédure utilisée consistait elle aussi à atteindre la correspondance entre le flou d'un contour fixe simulé à partir de mesures physiques et le flou de mouvement perçu sur un contour en mouvement. Cependant, l'ajustement était effectué en faisant varier la vitesse du contour en mouvement : c'est la principale différence avec notre expérience. Pour chacune des 8 transitions testées et pour les 4 vitesses,



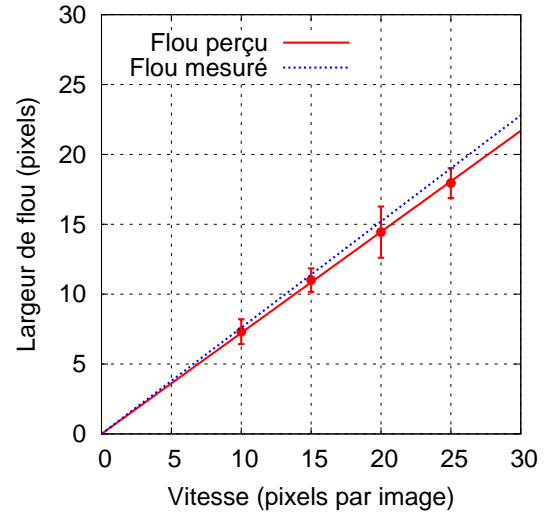
(a) Transition 0 → 63 sur l'écran 1.



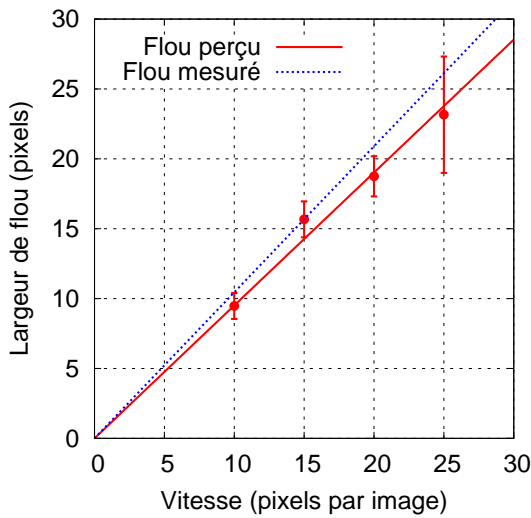
(b) Transition 63 → 191 sur l'écran 1.



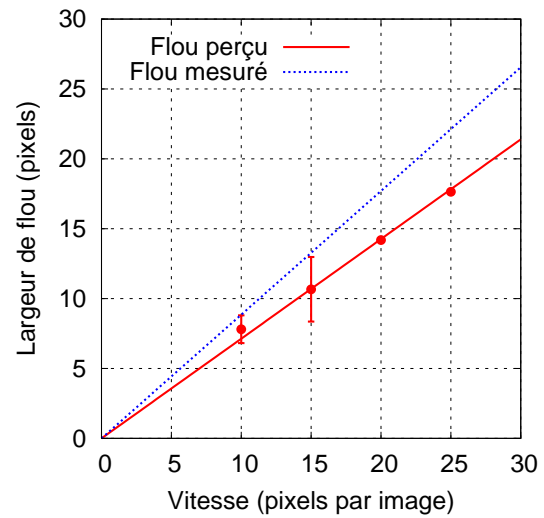
(c) Transition 0 → 63 sur l'écran 2.



(d) Transition 63 → 191 sur l'écran 2.



(e) Transition 0 → 63 sur l'écran 3.



(f) Transition 63 → 191 sur l'écran 3.

FIGURE 6.3 – Flou mesuré (ligne pointillée) et flou perçu (points) en fonction de la vitesse de déplacement du contour, pour les transitions 0 → 63 (à gauche) et 63 → 191 (à droite), sur chacun des trois écrans. Les barres verticales représentent l'écart-type des données subjectives.

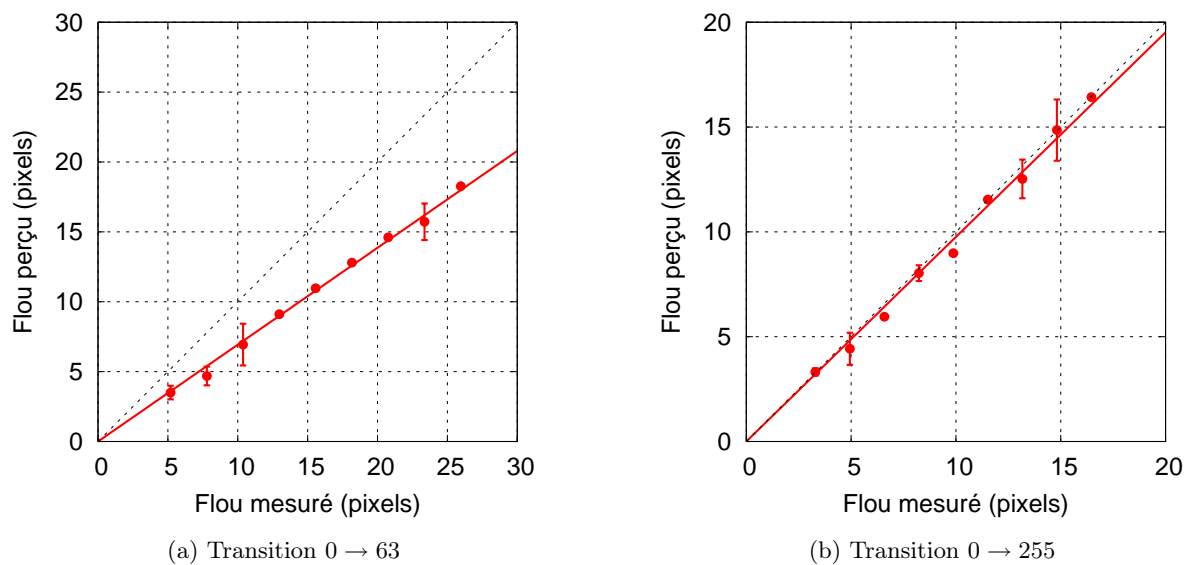


FIGURE 6.4 – Flou de mouvement perçu en fonction du flou de mouvement mesuré, pour les transitions $0 \rightarrow 63$ (a) et $0 \rightarrow 255$ (b), pour l'écran 1. Les barres verticales représentent l'écart-type des données subjectives. La droite correspond au modèle $BEW_p = A \cdot BEW_m$, avec $A = 0.69$ pour la transition $0 \rightarrow 63$ et $A = 0.98$ pour la transition $0 \rightarrow 255$.

leurs résultats indiquent que la vitesse ajustée est très proche de la vitesse utilisée lors des mesures physiques. Le gain entre le flou perçu et le flou mesuré était donc égal à 1 pour toutes les transitions. La différence de tâche entre nos deux expériences peut expliquer la divergence des résultats obtenus. Il convient également de noter que les stimuli utilisés ne sont pas tout à fait les mêmes. En effet, dans leurs expérimentations, Teunissen *et al.* ont utilisé un bloc de pixels de niveau de gris N_f en mouvement sur un fond uniforme de niveau de gris N_i . Par conséquent, la procédure d'égalisation de flou a été réalisée simultanément sur les transitions $N_i \rightarrow N_f$ et $N_f \rightarrow N_i$, alors que dans nos expériences, seule une transition était testée à la fois. L'égalisation simultanée de la transition montante et de la transition descendantes peut avoir introduit un biais dans les résultats obtenus par Teunissen *et al.*

6.4.3 Étude de la variation du coefficient A

Il peut être observé que le coefficient A varie en fonction de la transition considérée. Il est plus faible pour des contours de luminance faible, et tend vers 1 lorsque la luminance du contour augmente. Par exemple, la figure 6.4 donne le flou de mouvement perçu en fonction du flou de mouvement mesuré pour les transitions $0 \rightarrow 63$ et $0 \rightarrow 255$ sur l'écran 1. La droite correspond à l'optimisation de la relation 6.4 pour chacune des deux transitions. Le coefficient A correspond à la pente de cette droite, il est égal à 0.69 pour la transition $0 \rightarrow 63$ et à 0.98 pour la transition $0 \rightarrow 255$.

Pour étudier ce phénomène, nous considérons le profil spatial du contour en luminance perçue L^* . La luminance perçue est obtenue à partir de la luminance physique L en appliquant la relation A.2

N	Luminance perçue		
	écran 1	écran 2	écran 3
0	1.7	1.2	1.2
63	26	26.4	20.5
127	53.7	53.7	49.0
191	78.6	77.7	75.5
255	100	100	100

TABLE 6.2 – Luminance perçue L^* associée à chaque niveau de gris testé, pour chacun des trois écrans.

décrite en annexe :

$$L^* = \begin{cases} 116 \left(\frac{L}{L_{max}} \right)^{\frac{1}{3}} - 16 & \text{pour } \frac{L}{L_{max}} > 0.008856 \\ 903.3 \frac{L}{L_{max}} & \text{sinon} \end{cases}$$

Les luminances perçues associées aux niveaux de gris utilisés durant ces tests sont données dans le tableau 6.2. Les luminances perçues sont réparties équitablement entre 0 et 100 (unités arbitraires de l'espace de couleur CIE $L^*a^*b^*$). Les légères variations sont dues aux différences de fonctions gamma entre les écrans. À chaque transition $N_i \rightarrow N_f$ on associe donc deux valeurs de luminance perçue, la luminance perçue initiale L_i^* et la luminance perçue finale L_f^* .

La figure 6.5a présente l'évolution du coefficient A en fonction de la luminance perçue finale L_f^* sur l'écran 1, pour les transitions montantes ($N_i < N_f$). On constate que le coefficient A est proportionnel à la luminance perçue finale L_f^* . La luminance perçue de la partie finale du contour (à gauche du contour pour un mouvement de la gauche vers la droite comme ce fut le cas dans nos expérimentations) conditionne donc la perception du flou de mouvement sur ce contour. Plus la luminance finale est élevée, plus le coefficient de proportionnalité A entre flou perçu et flou mesuré s'approche de 1. À l'inverse lorsque la luminance finale est plus faible, le flou de mouvement perçu est plus faible que le flou mesuré. Pour les transitions descendantes ($N_f < N_i$), le comportement est inversé (figure 6.5b). Le coefficient A diminue à mesure que la luminance finale augmente. Lorsque la luminance finale du contour est faible, le flou de mouvement perçu est proche du flou de mouvement mesuré. Lorsque la luminance augmente, le flou de mouvement perçu est plus faible que le flou mesuré.

6.4.3.1 Transitions montantes

La figure 6.6 présente l'évolution du coefficient A en fonction de la luminance perçue finale L_f^* pour les écrans 1 et 2. Les quatre valeurs sont obtenues en calculant la moyenne sur l'ensemble des transitions de mêmes luminance finale, l'intervalle de confiance à 95% sur la moyenne est également

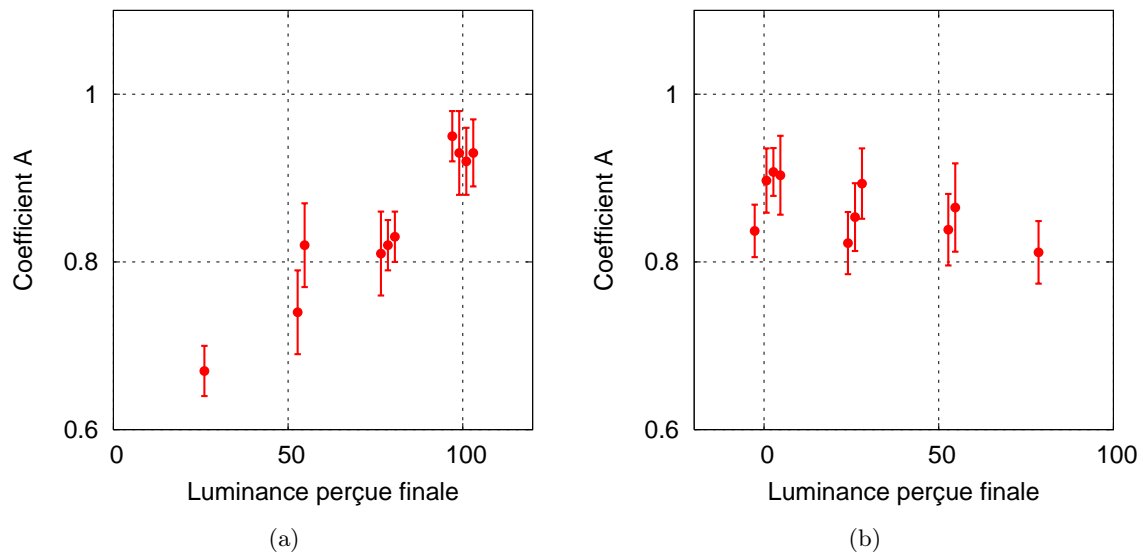


FIGURE 6.5 – Coefficient A entre le flou perçu et le flou mesuré sur l'écran 1, en fonction de la luminance perçue finale du contour L_f^* ; pour les transitions montantes ($L_f^* > L_i^*$) (a) et pour les transitions descendantes ($L_f^* < L_i^*$) (b). Les barres verticales correspondent à l'intervalle de confiance à 95% sur la moyenne. Le décalage horizontal entre les points a été ajouté pour plus de lisibilité.

représenté. Le coefficient de corrélation linéaire entre le coefficient A et la luminance perçue finale L_f^* est important, il vaut 0.9837 sur l'écran 1 et 0.9208 sur l'écran 2. Sur l'écran 3 (non représenté), aucune tendance particulière n'a été observée entre valeur de A et luminance du contour. Cela peut être dû au fait que l'écran 3 présente des artefacts de sur-commande sur sa réponse indicielle (cf. 3.3.2.3). Par conséquent, le profil spatial du contour en mouvement se trouve lui aussi affecté par ces artefacts et l'égalisation peut être plus difficile à atteindre.

La figure 6.7 présente les mêmes résultats mais avec cette fois la valeur moyenne du coefficient A sur l'ensemble des trois écrans testés. Ici encore, la moyenne a été réalisée sur l'ensemble des transitions montantes possédant la même luminance perçue finale L_f^* . L'intervalle de confiance à 95% est représenté (barres verticales) ainsi que la légère dispersion des luminances perçues finales d'un écran à l'autre (les barres horizontales représentent l'écart-type entre les 3 valeurs). La relation entre le coefficient A et la luminance perçue finale L_f^* sur l'ensemble des écrans testés est linéaire croissante, on a :

$$A_{mont} = 0.235 \frac{L_f^*}{100} + 0.713 \quad (6.6)$$

Le coefficient de corrélation linéaire entre ce modèle et les résultats expérimentaux est de 0.9934.

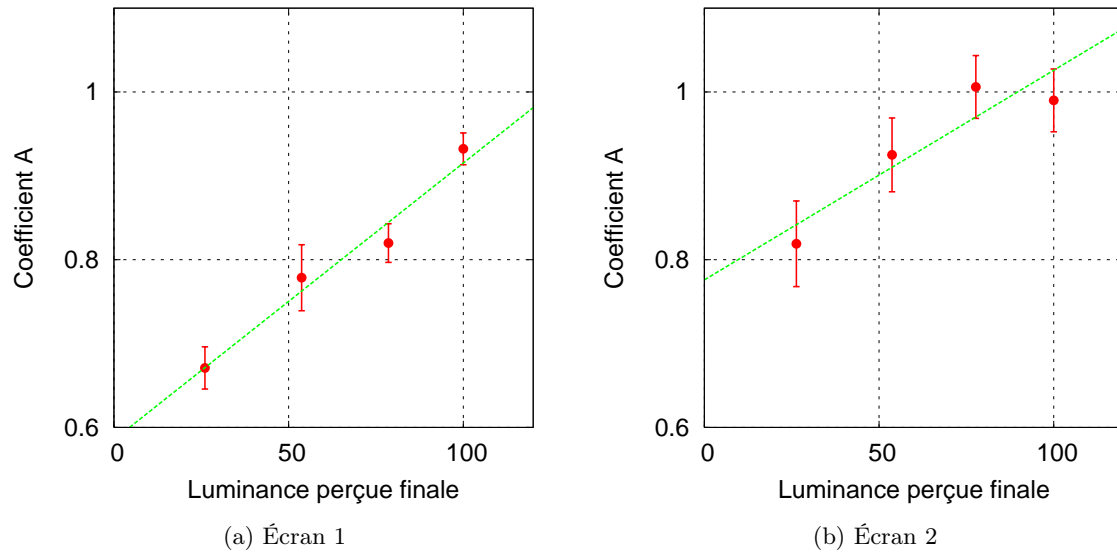


FIGURE 6.6 – Coefficient A entre le flou perçu et le flou mesuré, en fonction de la luminance perçue finale du contour L_f^* , pour les transitions montantes. Les valeurs sont obtenues en calculant la moyenne sur l'ensemble des transitions de même luminance finale. Les barres verticales représentent l'intervalle de confiance à 95% sur cette moyenne.

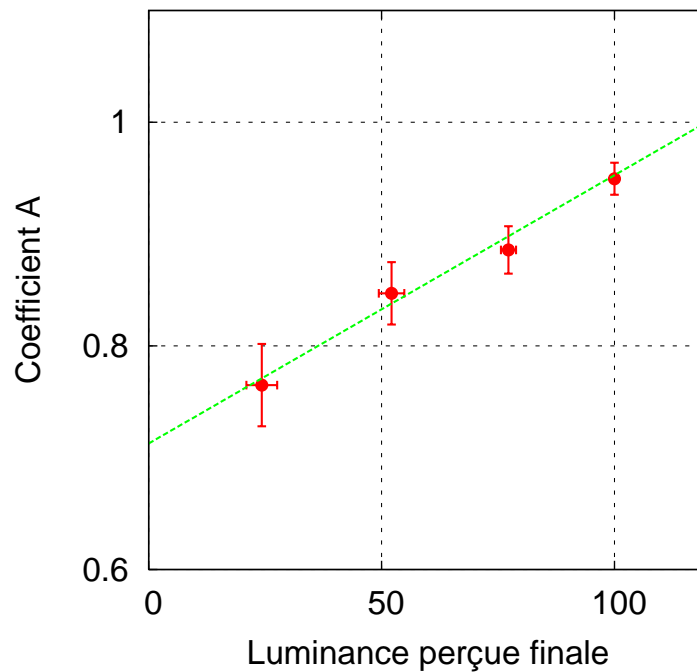


FIGURE 6.7 – Coefficient A entre le flou perçu et le flou mesuré, en fonction de la luminance perçue finale du contour L_f^* , pour les transitions montantes et pour les 3 écrans testés. Les barres horizontales correspondent à l'écart-type sur la valeur de la luminance perçue finale (cf. tableau 6.2).

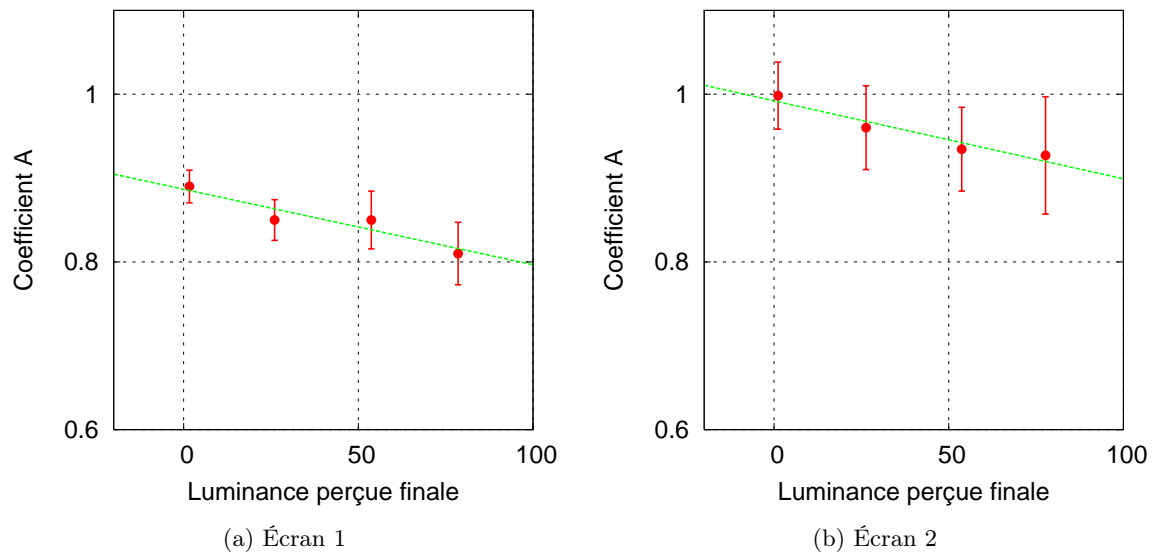


FIGURE 6.8 – Coefficient A entre le flou perçu et le flou mesuré, en fonction de la luminance perçue finale du contour L_f^* , pour les transitions descendantes. Les valeurs sont obtenues en calculant la moyenne sur l'ensemble des transitions de même luminance finale. Les barres verticales représentent l'intervalle de confiance à 95% sur cette moyenne.

6.4.3.2 Transitions descendantes

Les mêmes résultats ont été calculés pour les transitions descendantes. La figure 6.8 présente l'évolution du coefficient A en fonction de la luminance perçue finale L_f^* pour les écrans 1 et 2. Le coefficient de corrélation linéaire entre le coefficient A et la luminance perçue finale L_f^* est de 0.9477 sur l'écran 1 et de 0.9636 sur l'écran 2. Encore une fois, les résultats obtenus sur l'écran 3 ne permettent pas de tirer de conclusions individuelles. La figure 6.9 présente les résultats pour l'ensemble des trois écrans. On obtient également une relation linéaire entre le coefficient A et la luminance perçue finale L_f^* , mais elle est cette fois-ci décroissante, on a :

$$A_{desc} = -0.077 \frac{L_f^*}{100} + 0.924 \quad (6.7)$$

Le coefficient de corrélation linéaire entre ce modèle et les résultats expérimentaux est de 0.9839.

6.4.4 Évaluation globale du modèle

On cherche ici à obtenir une mesure du flou de mouvement qui soit un bon estimateur du flou de mouvement perçu. On vient de le voir, la mesure du flou de mouvement de manière photométrique donne un résultat éloigné du flou réellement perçu. Le coefficient de corrélation linéaire global entre

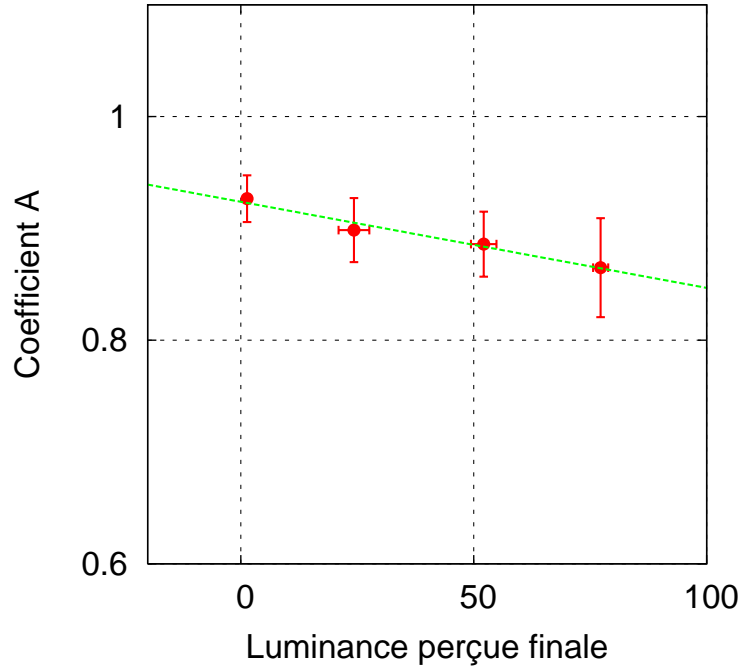


FIGURE 6.9 – Coefficient A entre le flou perçu et le flou mesuré, en fonction de la luminance perçue finale du contour L_f^* , pour les transitions descendantes et pour les 3 écrans testés. Les barres horizontales correspondent à l'écart-type sur la valeur de la luminance perçue finale (cf. tableau 6.2).

le temps de flou mesuré BET_m et le temps de flou perçu BET_p , sur les 60 transitions mesurées (20 transitions pour 3 écrans différents) est de 0.7947, l'erreur absolue moyenne est de 1.81 ms, et la racine carrée de l'erreur quadratique moyenne est de 2.4.

Les deux modèles définis précédemment permettent d'estimer la valeur du flou de mouvement perçu à partir du flou de mouvement mesuré et de la luminance du contour considéré. Le temps de flou perçu estimé BET_{est} est donné par :

$$BET_{est, N_i \rightarrow N_f} = \begin{cases} A_{mont}(L_f^*) BET_{m, N_i \rightarrow N_f} & \text{si } N_i < N_f \\ A_{desc}(L_f^*) BET_{m, N_i \rightarrow N_f} & \text{si } N_i > N_f \end{cases} \quad (6.8)$$

Le coefficient de corrélation linéaire entre le temps de flou perçu estimé BET_{est} donné par le modèle et le temps de flou perçu BET_p obtenu par la campagne de tests psychophysiques est de 0.8519, l'erreur absolue moyenne de 1.00 ms, et l'erreur quadratique moyenne de 1.43. Le modèle n'est pas parfait et, comme on a pu le voir, les résultats sont fortement dépendant de l'écran mesuré. De plus, ce modèle moyen est construit à partir des mesures effectuées sur seulement 3 écrans. Néanmoins, l'utilisation de ce modèle permet d'obtenir une estimation assez correcte du temps de flou perçu. La figure 6.10 illustre la relation entre le temps de flou perçu, le temps de flou mesuré et le temps de flou estimé.

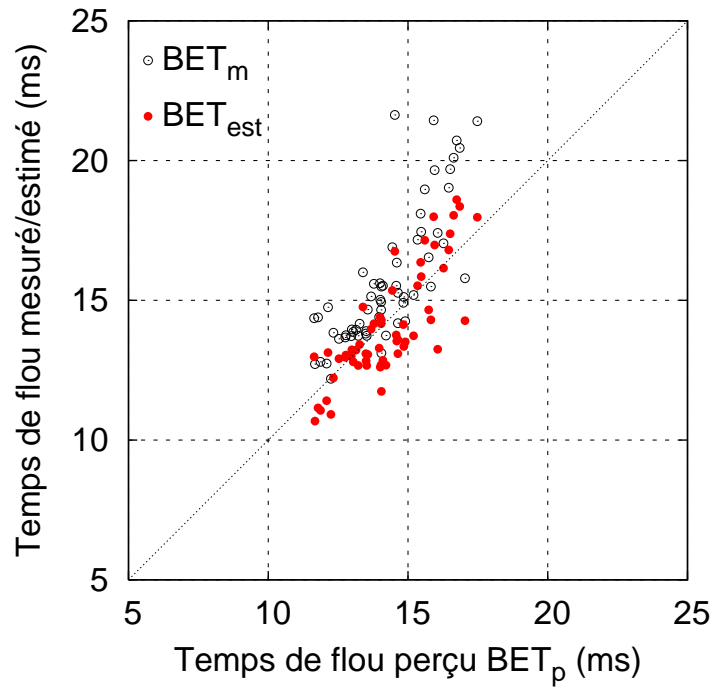


FIGURE 6.10 – Temps de flou mesuré BET_m et temps de flou estimé BET_{est} en fonction du temps de flou perçu BET_p , pour l'ensemble des 60 transitions mesurées.

6.5 Vers une mesure perceptive du flou de mouvement

Les tests psychophysiques présentés dans ce chapitre ont permis de mesurer précisément le temps flou perçu sur chacune des transitions testées et sur chacun des trois écrans. Pour chaque écran, nous disposons donc d'une valeur moyenne expérimentale du temps de flou perçu. Elle est notée MB_{EXP} et calculée comme la moyenne des valeurs de flou de mouvement perçu sur chaque transition, obtenues par l'expérience :

$$MB_{EXP} = \frac{1}{N} \sum_{N_i \rightarrow N_f} BET_{p, N_i \rightarrow N_f} \quad (6.9)$$

De plus, cette campagne de mesures psychophysiques a permis de montrer que le temps de flou perçu est globalement plus faible que le temps de flou mesuré de façon instrumentale. Cette atténuation du flou de mouvement perçu dépend de la luminance du contour en déplacement. Par conséquent, le temps de flou global MB_{VESA} , défini par le groupe de travail FPDM [VESA 05] (section 3.4.4) et calculé de la manière suivante :

$$MB_{VESA} = \frac{1}{N} \sum_{N_i \rightarrow N_f} BET_{m, N_i \rightarrow N_f},$$

ne constitue pas une bonne estimation du temps de flou perçu.

Pour tenir compte de la non-uniformité du système visuel humain dans la perception du flou de mouvement, une estimation est réalisée grâce à la pondération perceptuelle des valeurs de temps de

	Écran 1	Écran 2	Écran 3
MB_{EXP}	15.1	13.6	14.1
MB_{VESA}	17.8 (+18%)	14.1 (+4%)	15.3 (+9%)
MB_{PERC}	15.9 (+5%)	13.2 (-3%)	13.7 (-3%)

TABLE 6.3 – Valeur globale du temps de flou pour les 3 écrans testés (en millisecondes). MB_{VESA} est la moyenne des temps de flou mesuré, MB_{EXP} est la moyenne des temps de flou perçu obtenus par l'expérience, et MB_{PERC} est la moyenne des temps de flou estimés par le modèle décrit dans ce chapitre.

flou mesurées. Cette pondération est effectuée à partir des deux modèles définis précédemment pour l'ensemble des transitions montantes et pour l'ensemble des transitions descendantes (équations 6.6 et 6.7). L'estimation de la valeur moyenne du temps de flou perçu est notée MB_{PERC} et calculée de la façon suivante :

$$MB_{PERC} = \frac{1}{N} \sum_{N_i \rightarrow N_f} BET_{est, N_i \rightarrow N_f} \quad (6.10)$$

$$MB_{PERC} = \frac{1}{N} \left(\sum_{N_i < N_f} A_{mont}(L_f^*) BET_{m, N_i \rightarrow N_f} + \sum_{N_i > N_f} A_{desc}(L_f^*) BET_{m, N_i \rightarrow N_f} \right) \quad (6.11)$$

Cette pondération, mise en œuvre à partir des résultats des tests psychophysiques décrits dans ce chapitre, permet d'obtenir une valeur globale du temps de flou qui tient compte de la perception du flou de mouvement.

À titre d'exemple, le tableau 6.3 donne la valeur globale du temps de flou perçu MB_{EXP} , calculé à partir des résultats de l'expérience. Le temps de flou global MB_{VESA} , mesuré selon la méthode VESA, et l'estimation du temps de flou global perceptif MB_{PERC} calculée par une pondération des mesures photométriques, sont également donnés. On observe que le temps de flou global mesuré selon la méthode VESA est assez éloigné de la réalité, avec une erreur relative moyenne de l'ordre de 10%. Notre estimation du temps de flou perçu global MB_{PERC} est plus proche des valeurs obtenues par l'expérience, avec une erreur relative moyenne inférieure à 5%.

6.6 Conclusion

Les expérimentations mises en œuvre dans ce chapitre ont permis d'évaluer la perception du flou de mouvement sur les écrans à cristaux liquides. Les résultats obtenus montrent que le flou de mouvement perçu varie de façon linéaire avec la vitesse, mais avec un facteur de proportionnalité plus faible que celui observé pour le flou mesuré. Le coefficient de proportionnalité entre le flou perçu et le flou mesuré est donc inférieur à 1, indiquant qu'un contour en mouvement est perçu

moins flou qu'il ne l'est photométriquement. Ce phénomène est d'autant plus vrai que la luminance du contour est faible. Le rapport entre le temps de flou perçu et le temps de flou mesuré augmente avec la luminance du contour pour les transitions montantes. Il décroît avec la luminance pour les transitions descendantes. Cette tendance se retrouve d'un écran à l'autre mais dans des proportions différentes, elle est très peu marquée sur l'un des trois écrans testés.

La rapport entre le temps de flou perçu et le temps de flou mesuré peut donc être prédit à partir de la luminance du contour. Cette prédiction a permis la construction de deux relations linéaires qui permettent d'estimer le temps de flou perçu (résultats de l'expérience) à partir du temps de flou mesuré (résultats des mesures photométriques), pour chaque transition. Ces deux modèles concernent l'un les transitions montantes, l'autre les transitions descendantes. Ils consistent en une pondération des mesures photométriques du flou de mouvement par la luminance de la transition considéré, et permettent d'obtenir des valeurs de flou de mouvement perceptuellement uniformes. La moyenne de ces valeurs estimées sur l'ensemble des transitions constitue donc une estimation de la valeur globale du temps de flou perçu. Sur les trois écrans testés, l'écart moyen entre cette estimation et les résultats de l'expérience est de l'ordre de 5%. À titre de comparaison, l'écart moyen entre les résultats de l'expérience et la valeur globale recommandée par l'association VESA [VESA 05] (qui consiste en la moyenne des valeurs de flou de mouvement mesurées photométriquement, non pondérées) est supérieur à 10%. La moyenne des valeurs de flou de mouvement mesuré, pondérée par la luminance de chaque transition, constitue donc une valeur de flou perceptuellement uniforme, qui permet d'estimer globalement le flou perçu par les observateurs.

De manière à étudier plus profondément l'impact du flou de mouvement sur le ressenti global des observateurs, nous avons décidé de poursuivre les mesures subjectives avec des contenus plus complexes qu'un simple contour en mouvement. Dans le chapitre suivant nous nous focaliserons cette fois-ci sur la mesure de la gêne introduite par le flou de mouvement, en utilisant des séquences synthétiques aux mouvements contrôlés.

Chapitre 7

Gêne visuelle associée au flou de mouvement

Sommaire

7.1	Introduction	121
7.2	Expérimentations	123
7.2.1	Écrans et conditions de visualisation	123
7.2.2	Séquences	123
7.2.3	Procédure	125
7.2.4	Observateurs	125
7.3	Résultats	126
7.3.1	Détermination des vitesses seuil	126
7.3.2	Conversion des vitesses seuil	126
7.4	Observations et discussion	132
7.4.1	Valeurs de flou de mouvement aux vitesses seuil	132
7.4.2	Conséquences sur la valeur recommandée du temps de flou perçu global	135
7.5	Conclusion	137

7.1 Introduction

L'étude de la perception du flou de mouvement menée dans le chapitre précédent a montré qu'il était possible d'estimer le temps de flou perçu (calculé à partir des résultats des expérimentations psychophysiques) à partir du temps de flou mesuré (calculé à partir des mesures objectives du flou de mouvement). Cette estimation, réalisée à travers une pondération perceptuelle du temps de flou mesuré basée sur la luminance perçue des transitions considérées, permet de définir une métrique perceptuellement uniforme du flou de mouvement. La seconde étape de cette caractérisation psychovisuelle du flou de mouvement concerne l'évaluation de la gêne associée à la présence de flou dans les contenus affichés. En effet, il ne suffit pas de connaître la quantité de flou perçu, mais il convient également de connaître précisément l'impact de ce flou sur la qualité visuelle ressentie par

les observateurs. De cette manière, il sera possible d'émettre des recommandations concernant la métrique perceptive du flou de mouvement définie dans le chapitre précédent.

S'il existe de nombreuses études concernant la caractérisation photométrique du flou de mouvement, très peu de travaux de recherche ont cherché à évaluer la gêne associée au flou de mouvement sur les écrans à cristaux liquides. On peut cependant citer le travail de Yamamoto *et al.* [Yamamoto 06], dans lequel un ensemble d'observateurs a évalué l'altération de la qualité visuelle de deux séquences vidéo sur sept écrans LCD aux caractéristiques techniques différentes. Les résultats de cette évaluation subjective ont ensuite été comparés avec une mesure globale du flou de mouvement telle que celle présentée dans le chapitre 3 (cf. section 3.4.4). L'inconvénient de cette étude réside dans l'utilisation de séquences vidéo naturelles, ce qui interdit d'associer la gêne ressentie à une valeur précise de flou de mouvement et oblige la comparaison avec une valeur globale de flou. De plus, cette valeur globale de flou de mouvement est obtenue à partir de mesures photométriques et ne tient donc pas compte de la non-uniformité de la perception visuelle du flou de mouvement mise en évidence par les expériences du chapitre précédent. Au final, la corrélation entre le flou de mouvement mesuré et la gêne ressentie s'est avérée plutôt faible et le manque de conditions testées a conduit à une estimation plutôt difficile d'un seuil de tolérance au flou de mouvement. Les travaux de Shin *et al.* [Shin 06] méritent également qu'on s'y attarde. Dans cette étude, une évaluation subjective de la gêne ressentie a été menée avec un protocole similaire. Les stimuli utilisés furent différents. Il s'agissait ici d'images naturelles et de textes, défilants sur l'écran à vitesse constante. L'utilisation de stimuli plus simples a donc permis de comparer la gêne ressentie pour différentes vitesses de défilement avec le flou de mouvement associé aux différents stimuli. Cependant, un seul écran a été testé et comme pour l'étude précédente, ce sont des valeurs de flou mesurées de manière objective qui ont été utilisées. Néanmoins, cette méthodologie a permis de déterminer des seuils de perception et d'acceptabilité du flou mouvement pour l'écran testé.

Dans ce chapitre, nous nous inspirons de ces travaux pour mesurer précisément la gêne ressentie par les observateurs lorsque la qualité des stimuli affichés est altérée par le flou de mouvement. Les stimuli choisis doivent être assez simples pour que l'on puisse connaître précisément la quantité de flou de mouvement qui sera perçue sur leurs contours, mais ils doivent également être assez élaborés pour que les observateurs puissent juger de la gêne ressentie. Pour répondre à ces deux exigences, nous avons conçu trois types de stimuli : des séquences synthétiques en niveau de gris, des séquences constituées d'un texte défilant, et des images panoramiques en translation horizontale. Ces stimuli sont présentés aux observateurs avec différentes vitesses de déplacement sur les trois écrans testés dans le chapitre précédent. Le score de gêne mesuré pour chaque présentation est ensuite comparé aux valeurs de flou de mouvement perçu obtenues précédemment. Ainsi, la quantité de flou de mouvement perçu et la gêne associée à ce flou de mouvement pourront être mis en correspondance.

	$BET_{m, N_i \rightarrow N_f}$			$BET_{p, N_i \rightarrow N_f}$		
	min	moy	max	min	moy	max
Écran 1	13.6	17.8	21.6	12.5	15.0	17.5
Écran 2	12.2	14.1	15.8	11.6	13.6	17.0
Écran 3	13.7	15.3	17.5	12.1	14.2	16.3

TABLE 7.1 – Valeurs minimale, moyenne et maximale du temps de flou mesuré BET_m et du temps de flou perçu BET_p (en ms), sur les 20 transitions $N_i \rightarrow N_f$, pour chacun des trois écrans testés.

7.2 Expérimentations

7.2.1 Écrans et conditions de visualisation

Les écrans testés lors de cette campagne de tests subjectifs sont les mêmes écrans à cristaux liquides que ceux utilisés lors de l'étude de la perception du flou (cf. tableau 6.1). Les conditions de visualisation sont également les mêmes que dans les expériences du chapitre précédent. Pour chaque écran, nous disposons donc de la mesure psychovisuelle du flou de mouvement perçu sur 20 transitions gris-à-gris. Le temps de flou mesuré BET_m et le temps de flou perçu BET_p , pour les trois écrans, sont donnés dans le tableau 7.1.

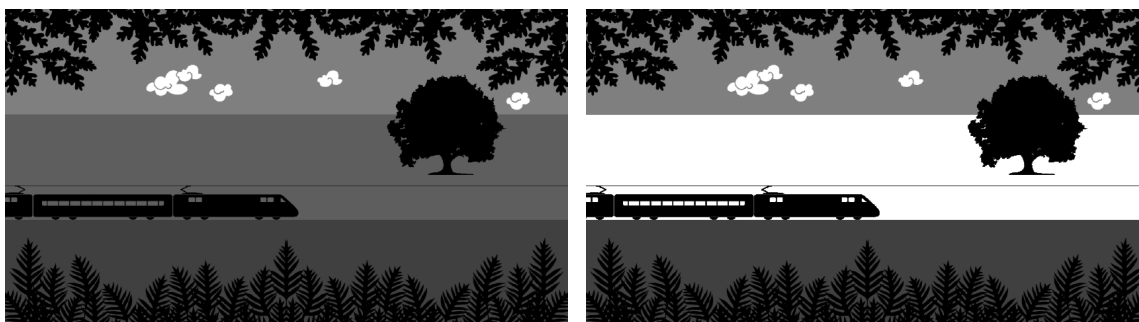
Comme précédemment, les expérimentations ont été réalisées conjointement par l'équipe IVC du laboratoire IRCCyN et par le Video and Display Quality Lab du centre de recherche Acreo AB en Suède. Deux salles de tests ont été mises à contribution, les écrans 1 et 3 ayant été testés à l'IRCCyN-IVC, et l'écran 2 chez Acreo.

7.2.2 Séquences

Un ensemble de 24 stimuli a été créé. Ces stimuli sont classés en trois catégories :

- Stimuli « Train » : 10 stimuli représentant un train se déplaçant de la gauche vers la droite. Les niveaux de gris du train et de l'arrière-plan sont différents d'un stimulus à l'autre. La hauteur du train (et donc des contours sur lesquels le flou apparaît) a été fixée à 2 degrés visuels.
- Stimuli « Texte » : 10 stimuli représentant un texte défilant de la droite vers la gauche. Les niveaux de gris du texte et de l'arrière-plan sont différents d'un stimulus à l'autre, l'arrière-plan peut également être une image naturelle. Le texte est en majuscule et sa taille a été fixée à 0.4 degré de hauteur, c'est à dire la taille moyenne des textes défilants rencontrés sur les chaînes télévisées d'informations (pour une distance d'observation standard).
- Stimuli « Panorama » : 4 images panoramiques représentant des scènes naturelles photographiées avec un angle de vue de 360°. Ces images peuvent donc être animées d'un mouvement de translation perpétuel.

La figure 7.1 donne deux illustrations de chacune des catégories. Pour ces 24 stimuli, l'objet d'intérêt (*i.e.* le train, le texte, ou l'image entière selon le type de stimulus considéré) est animé d'une vitesse constante. Huit vitesses ont été testées : 2, 4, 6, 8, 10, 12, 14 et 16 pixels par image (respectivement 3, 6, 9, 12, 15, 18 et 21 degrés par seconde à la distance d'observation considérée). Au final, ce sont



(a) Exemples de séquences « Train »



(b) Exemple de séquences « Texte »



(c) Exemple de séquences « Panorama »

FIGURE 7.1 – Exemples de séquences utilisées durant les expérimentations, pour chacun des trois types de contenus : séquences « Train » (a), séquences « Texte » (b) et séquences « Panorama » (c).

donc 192 séquences qui ont été évaluées par chacun des observateurs. La campagne de tests a été séparée en deux sessions de 96 séquences. La durée moyenne de chaque session était d'environ 20 minutes.

5	Imperceptible
4	Perceptible mais non gênant
3	Légèrement gênant
2	Gênant
1	Très gênant

TABLE 7.2 – Échelle de dégradations à cinq niveaux.

7.2.3 Procédure

Pour chaque session, les observateurs évaluent 96 séquences : 40 séquences « Train », 40 séquences « Texte » et 16 séquences « Panorama ». Ces 96 séquences sont présentées dans un ordre aléatoire différent pour chaque observateur. Les observateurs ont la possibilité de rejouer la séquence à tout moment, autant de fois qu'ils le souhaitent.

Après la visualisation d'une séquence, deux écrans de vote sont présentés à l'observateur successivement. Tout d'abord, l'observateur doit juger si le flou de mouvement présent dans la séquence était acceptable ou non (choix binaire). Ensuite, l'observateur doit juger dans quelle mesure le flou de mouvement a altéré la qualité globale de la séquence. Ce second jugement est effectué par l'intermédiaire d'une échelle de dégradations à cinq niveaux (cf. tableau 7.2).

7.2.4 Observateurs

Les écrans 1 et 3 ont été évalués au sein du laboratoire IVC-IRCCyN. 22 observateurs ont évalué les deux écrans, 10 observateurs ont évalué seulement l'écran 1 et 8 observateurs ont évalué seulement l'écran 3. Au final, il y a donc eu 32 observateurs pour l'écran 1 et 30 observateurs pour l'écran 3. Parmi les 22 observateurs à avoir réalisé le test sur les deux écrans, une moitié a commencé par l'écran 1, l'autre moitié par l'écran 3. L'âge des 40 observateurs impliqués varie de 18 à 42 ans, avec une moyenne de 26 ans. 24% d'entre eux sont de sexe féminin. L'écran 2 a été testé au sein du laboratoire Acreo AB. 32 observateurs ont participé à l'expérience. L'âge des observateurs varie de 12 à 67 ans, avec une moyenne de 38 ans. 59% d'entre eux sont de sexe féminin.

La vision de tous les observateurs a été testée de manière à vérifier qu'ils possédaient une acuité visuelle supérieure ou égale à 9/10. Tous furent familiarisés avec la procédure mise en œuvre dans les expériences grâce à une session d'entraînement comprenant plusieurs séquences et dont les résultats n'ont pas été pris en compte. Une fois l'ensemble des tests terminés, un procédé de rejet des observateurs incohérents a été effectué. Pour un stimulus donné, le coefficient de corrélation linéaire entre les huit résultats (un résultat par vitesse testée) de chaque observateur et les 8 résultats moyens (moyenne calculée sur l'ensemble des observateurs non rejetés) est calculé. L'observateur est rejeté si ce coefficient est inférieur à 0.8. Les valeurs moyennes sont alors recalculées avec les observateurs restants et le procédé est répété jusqu'à ce que les résultats de tous les observateurs soient corrélés avec les résultats moyens avec un coefficient supérieur à 0.8. Ce processus de rejet est effectué indépendamment pour chaque stimulus. En moyenne, 3.3 observateurs ont été rejetés

pour les stimuli « Train », 1.5 pour les stimuli « Texte » et 1.7 pour les stimuli « Panorama ». Dans tous les cas, plus de 25 observateurs ont été retenus pour chaque stimulus.

7.3 Résultats

7.3.1 Détermination des vitesses seuil

Pour chacun des 24 stimuli, le taux d'acceptabilité et le score de gêne moyen sont obtenus pour les huit vitesses testées. Un exemple de résultats obtenus est présenté figure 7.2, avec les intervalles de confiance à 95%. Pour chaque stimulus, ces résultats permettent de tracer deux fonctions, à partir desquelles on peut calculer trois vitesses seuil :

- Vitesse au seuil de perception (V_{perc}) : c'est la vitesse pour laquelle le score de gêne est égal à 4.5 (le flou de mouvement devient perceptible mais non gênant).
- Vitesse au seuil de gêne ($V_{gêne}$) : c'est la vitesse pour laquelle le score de gêne est égal à 3.5 (le flou de mouvement devient légèrement gênant).
- Vitesse au seuil de non acceptabilité (V_{accept}) : c'est la vitesse pour laquelle le taux d'acceptabilité est égale à 0.5 (50% des observateurs considère le flou de mouvement comme non acceptable).

Ces trois vitesses seuil ont été calculées par l'intermédiaire d'un ajustement de courbe sur les données mesurées. Elles sont représentées figure 7.2 par les points rouges situés sur les fonctions d'ajustement. Une fonction logistique a été utilisée pour l'acceptabilité et une fonction spline de degré 3 pour le score de gêne. L'intervalle de confiance de la vitesse au seuil de non acceptabilité V_{accept} est obtenu directement par le résultat de l'ajustement de courbe puisque cette valeur est un des paramètres de la fonction logistique. Pour les vitesses seuil V_{perc} et $V_{gêne}$, un ajustement de courbe a été réalisé sur les valeurs moyennes plus ou moins l'intervalle de confiance, de manière à obtenir des bornes sur la valeur des seuils. Ces bornes sont représentées par des triangles sur la figure 7.2.

7.3.2 Conversion des vitesses seuil

Les vitesses seuil déterminées pour chacun des 24 stimuli sont exprimées en pixels par image. Elles représentent les vitesses pour lesquelles le flou de mouvement perçu sur chaque stimulus devient respectivement perceptible, gênant et non acceptable. Bien entendu, à chaque vitesse correspond une largeur de flou de mouvement perçu qui dépend des transitions gris-à-gris mises en jeu. Pour chaque vitesse seuil, on peut donc calculer la valeur seuil du flou de mouvement perçu en utilisant la relation 6.3 déterminée dans le chapitre précédent pour chaque transition $N_i \rightarrow N_f$.

Le flou de mouvement perçu BEW_p correspondant à la vitesse de déplacement déterminée au seuil de perception est noté BEW_p^{perc} . Il est calculé de la façon suivante :

$$BEW_p^{perc} = BET_p \cdot V_{perc} \quad (7.1)$$

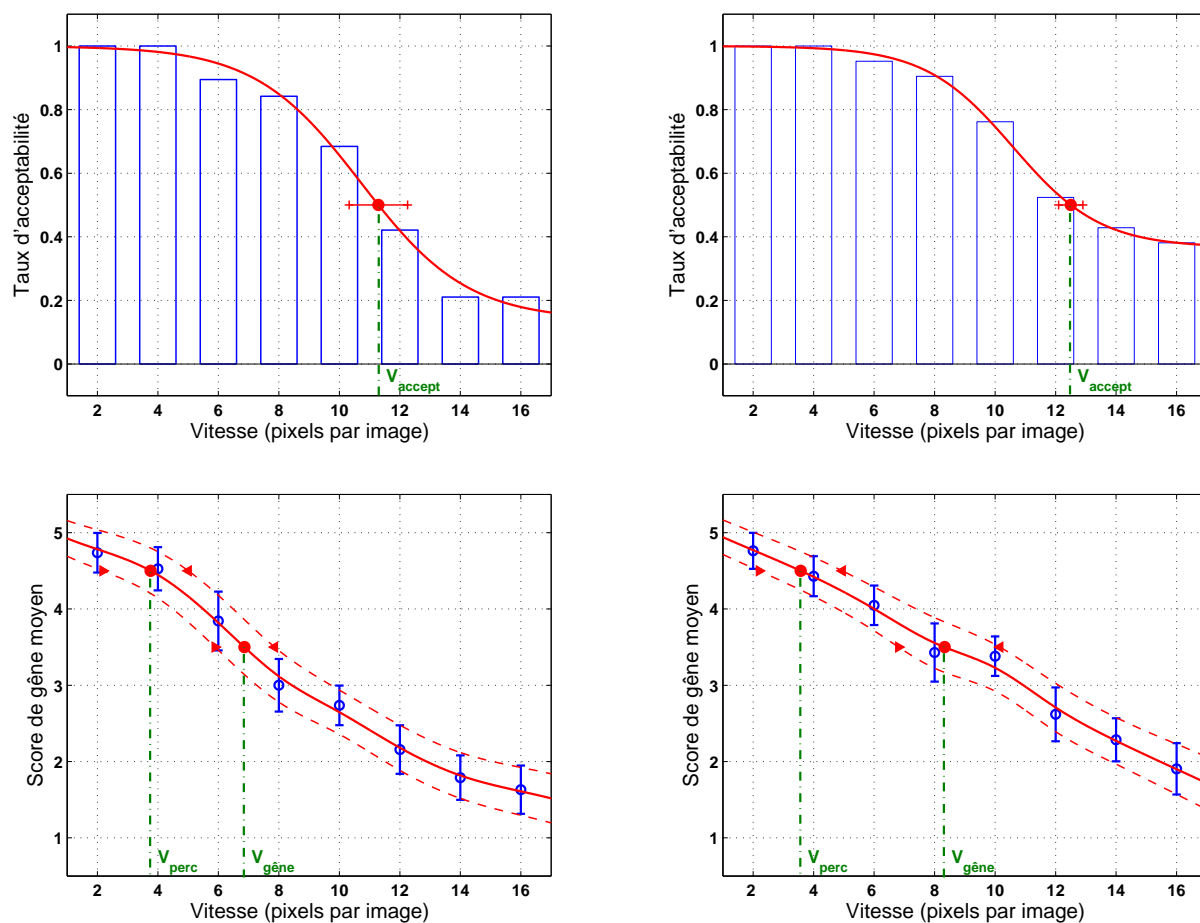


FIGURE 7.2 – Exemples de résultats obtenus sur l'écran 2 pour deux stimuli. Le taux d'acceptabilité (en haut) et le score de gène moyen (en bas) sont donnés en fonction de la vitesse du mouvement. Les deux stimuli présentés sont la séquence « Panorama » *Seine* (à gauche) et la séquence « Train » avec une transition $127 \rightarrow 255$ (à droite).

De la même façon, le flou de mouvement perçu à la vitesse déterminée au seuil de gêne est noté $BEW_p^{gêne}$ et le flou de mouvement perçu à la vitesse déterminée au seuil de non acceptabilité est noté BEW_p^{accept} :

$$\begin{aligned} BEW_p^{gêne} &= BET_p \cdot V_{gêne} \\ BEW_p^{accept} &= BET_p \cdot V_{accept} \end{aligned}$$

La valeur du temps de flou perçu BET_p utilisée dans les trois équation ci-dessus dépend de la ou des transitions gris-à-gris qui sont mises en œuvre dans la séquence considérée. Dans la suite de cette section, la méthode utilisée pour déterminer ce temps de flou perçu est précisée pour chaque type de séquence et la largeur de flou de mouvement perçu aux trois seuils est donnée pour chaque stimuli.

7.3.2.1 Séquences « Train »

Le tableau 7.3 présente les vitesses mesurées au seuil de perception V_{perc} pour chacune des dix séquences de type « Train », sur l'écran 1. Compte tenu de la configuration de ce type de séquence, les largeurs de flou perçu associées à ces vitesses seuil sont calculées en se basant sur les résultats des expériences psychovisuelles obtenus avec la transition gris-à-gris $N_i \rightarrow N_f$, où N_i est le niveau de gris de l'arrière plan, et N_f le niveau de gris du train. Pour chaque séquence, le temps de flou perçu $BET_{p, N_i \rightarrow N_f}$ (calculé dans le chapitre précédent) est donc utilisé pour obtenir la largeur de flou perçu au seuil de perception $BEW_{p, N_i \rightarrow N_f}^{perc}$. La moyenne et l'écart-type de la largeur de flou perçu sur l'ensemble des dix séquences de type « Train » sont également indiqués dans le tableau. Les largeurs de flou perçu sont également calculées, de la même façon, aux seuils de gêne et au seuil de non acceptabilité, les résultats sont présentés dans les tableaux 7.4 et 7.5.

Pour les trois vitesses seuil considérées, nous avons également calculé les largeurs de flou mesuré au seuil de perception BEW_m^{perc} . La même méthode que celle décrite ci-dessus a été employée, mais en utilisant le temps de flou mesuré BET_m au lieu du temps de flou perçu BET_p . Ces largeurs de flou mesuré sont également présentées dans les tableaux 7.3, 7.4 et 7.5. On remarque (tableau 7.3) que les valeurs de flou mesuré BEW_m^{perc} sont assez dispersées sur l'ensemble des 10 séquences, avec un écart-type supérieur à 0.6 (18% de la moyenne). A l'inverse, les valeurs de flou de mouvement perçu BEW_p^{perc} , obtenues à partir des mesures psychophysiques décrites dans le chapitre précédent, sont beaucoup moins dispersées, avec un écart-type de 0.29 (10% de la moyenne). Ce résultat confirme que les largeurs de flou obtenues via les expérimentations de mesure de la perception du flou correspondent bien à des valeurs de flou de mouvement *perçu*, et donc indépendantes de la transition mesurée (perceptuellement uniforme). Les mêmes remarques s'appliquent pour le seuil de gêne (tableau 7.4), pour lequel la dispersion du flou de mouvement mesuré est de 1.66 (24% de la moyenne), tandis qu'elle n'est que de 0.95 (16% de la moyenne) pour le flou de mouvement perçu, ainsi que pour le seuil de non acceptabilité (tableau 7.5), où la dispersion du flou mesuré est de 2.08 (17% de la moyenne) et celle du flou perçu de 0.64 (6% de la moyenne).

Séquence	N_i	N_f	V_{perc}	BEW_m^{perc}	BEW_p^{perc}
1	0	63	3.6	4.7	3.1
2	0	127	3.1	3.9	2.9
3	0	255	3.5	2.9	2.7
4	63	0	3.8	3.7	3.1
5	127	0	3.1	2.9	2.6
6	127	255	3.7	3.1	2.8
7	191	255	3.7	3.1	2.9
8	255	0	3.1	2.8	2.6
9	255	127	3.3	3.8	3.3
10	255	191	3.6	3.0	3.5
Moyenne			3.5	3.4	3.0
Écart-type			0.29	0.61	0.29

TABLE 7.3 – Vitesse au seuil de perception du flou de mouvement V_{perc} (en pixels par image), et largeurs de flou de mouvement mesuré BEW_m^{perc} et perçu BEW_p^{perc} correspondantes (en pixels), pour les séquences « Train », sur l'écran 1. N_i est le niveau de gris de l'arrière-plan, et N_f est le niveau de gris du train.

Séquence	N_i	N_f	$V_{gêne}$	$BEW_m^{gêne}$	$BEW_p^{gêne}$
1	0	63	7.9	10.3	6.9
2	0	127	6.9	8.8	6.6
3	0	255	7.3	6.0	5.7
4	63	0	6.8	6.5	5.5
5	127	0	6.0	5.6	5.0
6	127	255	7.2	5.9	5.4
7	191	255	7.8	6.5	6.1
8	255	0	6.1	5.6	5.0
9	255	127	7.8	8.9	7.7
10	255	191	7.6	6.4	7.3
Moyenne			7.1	7.0	6.1
Écart-type			0.69	1.66	0.95

TABLE 7.4 – Vitesse au seuil de gêne du flou de mouvement $V_{gêne}$ (en pixels par image), et largeurs de flou de mouvement mesuré $BEW_m^{gêne}$ et perçu $BEW_p^{gêne}$ correspondantes (en pixels), pour les séquences « Train », sur l'écran 1. N_i est le niveau de gris de l'arrière-plan, et N_f est le niveau de gris du train.

Les mêmes opérations sont réalisées pour les deux autres écrans. Le tableau 7.6 donne les valeurs moyennes des largeurs de flou de mouvement perçu BEW_p correspondant à chaque seuil, pour chaque écran, ainsi que les valeurs moyennes sur l'ensemble des trois écrans. Les largeurs de flou

Séquence	N_i	N_f	V_{accept}	BEW_m^{accept}	BEW_p^{accept}
1	0	63	13.2	17.1	11.5
2	0	127	11.2	14.4	10.7
3	0	255	12.6	10.4	9.8
4	63	0	12.2	11.7	9.8
5	127	0	12.3	11.5	10.3
6	127	255	15.4	12.6	11.6
7	191	255	13.1	11.0	10.2
8	255	0	12.6	11.4	10.3
9	255	127	10.5	12.0	10.4
10	255	191	-	-	-
Moyenne			12.8	12.5	10.5
Écart-type			1.51	2.08	0.64

TABLE 7.5 – Vitesse au seuil de non acceptabilité du flou de mouvement V_{accept} (en pixels par image), et largeurs de flou de mouvement mesuré BEW_m^{accept} et perçu BEW_p^{accept} correspondantes (en pixels), pour les séquences « Train », sur l'écran 1. N_i est le niveau de gris de l'arrière-plan, et N_f est le niveau de gris du train. Le seuil de non acceptabilité n'a pas pu être mesuré pour la 10^{ème} transition (taux d'acceptabilité supérieur à 0.5 pour toutes les vitesses testées).

	Écran 1	Écran 2	Écran 3	Moyenne
BEW_p^{perc}	4.4	3.9	4.1	4.1
$BEW_p^{gêne}$	9.2	8.1	8.8	8.7
BEW_p^{accept}	15.8	11.8	15.7	14.4

TABLE 7.6 – Valeurs moyennes du flou de mouvement perçu BEW_p au seuil de perception, au seuil de gêne et au seuil de non acceptabilité, pour chaque écran, pour les séquences de type « Train ». Les largeurs de flou sont exprimées en minutes d'arc.

sont exprimées en minutes d'arc (1 pixel équivaut à 1.5 minutes d'arc à la distance d'observation fixée durant les tests). Pour les séquences de type « Train », l'amplitude moyenne du flou perçu au seuil de perception est de 4.1 minutes d'arc. Celle au seuil de gêne est de 9 minutes d'arc. Enfin l'amplitude de flou perçu au delà de laquelle celui-ci n'est plus acceptable est de 14.4 minutes d'arc.

7.3.2.2 Séquences « Texte »

On distingue deux types de séquences « Texte ». Une première partie consiste en un texte d'un niveau de gris N_f défilant sur un fond uniforme de niveau de gris N_i . Le flou perçu est alors calculé en prenant comme valeur de temps de flou perçu BET_p la moyenne des deux temps de flou perçu obtenus pour les transitions $N_i \rightarrow N_f$ et $N_f \rightarrow N_i$. En effet, on considère que la transition montante et la transition descendante, pour un couple (N_i, N_f) donné, sont toutes les deux impliquées dans la perception du flou de mouvement. Le tableau 7.7a donne les valeurs moyennes de largeur de flou perçu à chaque vitesse seuil et pour chaque écran. On remarque que ces valeurs sont très proches d'un écran à l'autre. Elles sont en moyenne plus faibles que celles obtenues pour les séquences de

	Écran 1	Écran 2	Écran 3	Moyenne
BEW_p^{perc}	3.6	3.5	3.9	3.7
$BEW_p^{gêne}$	6.4	6.5	6.7	6.5
BEW_p^{accept}	9.1	8.6	9.0	8.9

(a) Texte défilant sur un fond uniforme.

	Écran 1	Écran 2	Écran 3	Moyenne
BEW_p^{perc}	4.2	3.6	4.1	4.0
$BEW_p^{gêne}$	7.3	6.6	7.1	7.0
BEW_p^{accept}	9.9	8.8	9.6	9.4

(b) Texte défilant sur une image naturelle.

TABLE 7.7 – Valeurs moyennes du flou de mouvement perçu BEW_p aux seuils de perception, de gêne et de non acceptabilité, pour chaque écran, pour les séquences de type « Texte ». Les valeurs de flou sont exprimées en minutes d’arc. Les résultats sont séparés entre les séquences constituées d’un texte défilant sur un fond uniforme (a) et celles constituées d’un texte défilant sur une image naturelle (b).

type « Train ». On peut penser que la lisibilité du texte défilant plus que le flou perçu (même si ces deux paramètres sont liés) a été prise en compte par les observateurs pour construire leur jugement. Ce qui expliquerait pourquoi les valeurs seuil obtenues sont plus faibles.

La seconde partie des séquences consiste en un texte de niveau de gris N_f défilant sur une image naturelle. Le temps de flou considéré pour convertir la vitesse seuil en largeur de flou correspond dans ce cas à la moyenne des temps de flou sur toutes les transitions impliquant le niveau de gris N_f du texte. Le tableau 7.7b donne les valeurs moyennes de largeur de flou perçu à chaque vitesse seuil et pour chaque écran. Les valeurs de flou de mouvement perçu correspondant aux vitesses seuil sont plus élevées que pour les séquences constituées de texte défilant sur un fond uniforme (tableau 7.7a). La différence est de l’ordre d’une demi minute d’arc. Cela semble indiquer que le flou est moins perceptible et moins gênant dans les séquences texturées.

7.3.2.3 Séquences « Panorama »

Enfin, les séquences « Panorama » sont constituées d’une image naturelle animée d’un mouvement de translation horizontale. Chaque image contient un nombre important de contours, et toutes les transitions possibles sont susceptibles d’être présentes. Par conséquent, la largeur de flou perçu pour chaque vitesse seuil est calculée en utilisant la moyenne de l’ensemble des temps de flou perçu BET_p obtenus pour les vingt transitions étudiées (cf. relation 6.9). Le tableau 7.8 donne les valeurs moyennes des largeurs de flou perçu à chaque vitesse seuil et pour chaque écran, pour les séquences de type « Panorama ».

L’amplitude du flou perçu au seuil de perception est de l’ordre de 4 minutes d’arc, comme sur les deux autres types de séquences. Concernant le seuil de gêne et le seuil de non acceptabilité, les

	Écran 1	Écran 2	Écran 3	Moyenne
BEW_p^{perc}	4.3	3.7	4.4	4.0
$BEW_p^{gêne}$	7.8	7.1	8.1	7.7
BEW_p^{accept}	12.5	10.3	12.4	11.7

TABLE 7.8 – Valeurs moyennes du flou de mouvement perçu BEW_p aux seuils de perception, de gêne et de non acceptabilité pour chaque écran, pour les séquences de type « Panorama ». Les valeurs de flou sont exprimées en minutes d’arc.

valeurs obtenues avec les séquences de type « Panorama » sont entre les valeurs obtenues avec les séquences « Train » (plus élevées) et celles obtenues avec les séquences « Texte » (plus faibles).

7.4 Observations et discussion

7.4.1 Valeurs de flou de mouvement aux vitesses seuil

La figure 7.3 donne les valeurs moyennes des largeurs de flou de mouvement perçu aux vitesses seuil. Les résultats sont présentés pour chaque écran et pour chaque type de séquence. Ils sont disposés de deux façons différentes pour faciliter la comparaison entre les différentes conditions de test. Les intervalles de confiance à 95% sont également indiqués sur la figure. Il s’agit de la valeur moyenne des intervalles de confiance obtenus pour chaque séquence d’un type donné.

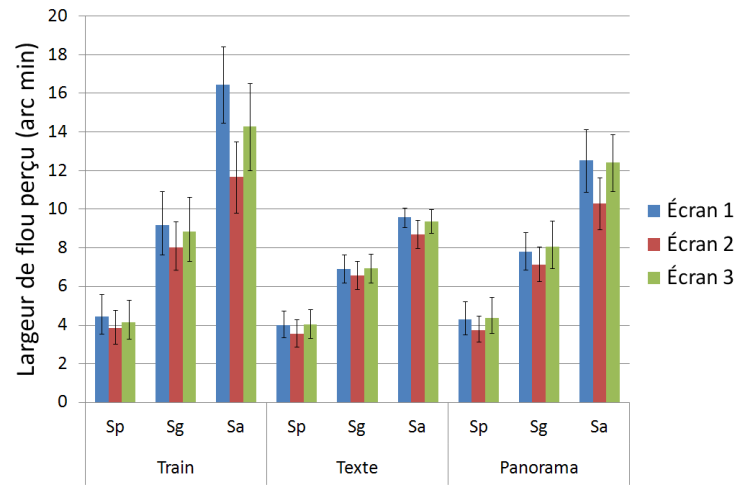
7.4.1.1 Seuil de perception

On remarque sur la figure 7.3b que la largeur de flou perçu au seuil de perception est très stable d’un écran à l’autre et d’un type de séquence à l’autre. Au vu des intervalles de confiance à 95%, il n’y a pas de différences significatives entre les différentes conditions de test. Ce résultat permet encore une fois de constater que les mesures subjectives du flou de mouvement perçu réalisées dans le chapitre 6 ont donné des résultats probants. En effet des variations sont observées sur la valeur de flou mesurée au seuil de perception, mais l’utilisation des temps de flou perçu BET_p permet d’obtenir une valeur quasiment constante du flou de mouvement perçu.

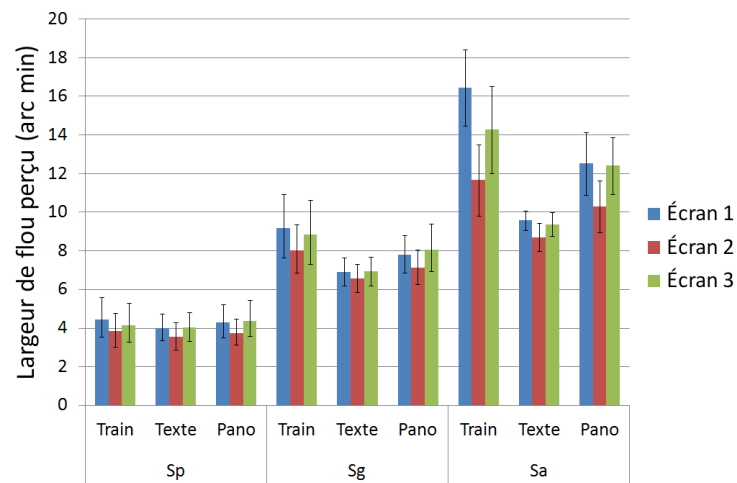
La valeur moyenne de la largeur de flou perçu au seuil de perception est égale à 4.0 minutes d’arc. L’écart-type sur les 9 valeurs moyennes est de 0.30.

7.4.1.2 Seuil de gêne

Le flou de mouvement perçu moyen au seuil de gêne est de 7.7 minutes d’arc, avec un écart-type sur les 9 valeurs moyennes de 0.91. Aucune différence notable n’est observée d’un écran à l’autre. La valeur moyenne est cependant légèrement plus importante pour les séquences de type « Train », ce qui indique que le flou de mouvement est moins gênant sur ces séquences synthétiques. Une des



(a)



(b)

FIGURE 7.3 – Valeur moyenne du flou de mouvement perçu au seuil de perception BEW_p^{perc} , au seuil de gêne BEW_p^{gene} et au seuil de non acceptabilité BEW_p^{accept} , pour chaque écran et pour chaque type de séquence. Les barres verticales figurent la moyenne des intervalles de confiance à 95%. Les deux sous-figures présentent les mêmes résultats mais groupés de façon différente.

explications de cette tendance tient au fait que, pour certains observateurs, le flou de mouvement ajoute du réalisme à ce type de séquences synthétiques et renforce l'impression de vitesse. Ceci peut expliquer pourquoi la tolérance vis à vis de cet artefact est plus élevée pour ce type de séquences.

7.4.1.3 Seuil de non acceptabilité

Contrairement aux deux valeurs seuil précédentes, la valeur du flou de mouvement perçu au seuil de non acceptabilité est sujette à variation. On remarque sur la figure 7.3a que le seuil de non acceptabilité est différent d'un type de séquence à l'autre. Il est assez élevé pour les séquences de types « Train », certainement pour les mêmes raisons que celles exprimées précédemment. La tolérance est plus faible pour les séquences de type « Panorama » et encore plus faible pour les séquences de type « Texte ». Dans le premier cas, on peut penser que la décision des observateurs est liée à la perte de résolution liée au flou de mouvement. Les images naturelles contiennent de nombreuses zones texturées et il paraît évident que le flou de mouvement global est plus gênant sur ce type de contenus. Pour le texte, la réponse des observateurs semble être liée à la lisibilité du texte. Celle-ci est bien entendu conditionnée par la taille du texte qui a été fixée à 24 minutes d'arc de hauteur (taille moyenne observée sur les chaînes d'informations pour une distance d'observation conforme aux recommandations). Le seuil de non acceptabilité pour un texte défilant est situé à une largeur de flou perçu de 9.2 minutes d'arc en moyenne, c'est à dire environ 40% de la taille du texte.

On remarque également une forte variation du seuil de non acceptabilité en fonction de l'écran testé. Plus précisément, cette variation semble être le fait de la différence des groupes d'observateurs. En effet, pour les écrans 1 et 3, dont les groupes d'observateurs respectifs sont communs à plus de 68%, les différences sont négligeables au vu de la valeur des intervalles de confiance. Sur l'écran 2, testé en Suède par un groupe d'observateurs différents, le seuil de non acceptabilité est globalement plus faible. Plusieurs facteurs peuvent expliquer cette différence. Les conditions d'observations ont pu être légèrement différentes d'un laboratoire à l'autre, mais la proximité des seuils de gêne et des seuils de perception semblent laisser penser que ces différences n'ont pas une influence importante. De même, il peut y avoir une certaine différence culturelle entre les observateurs suédois et français, bien que ce ne semble pas être pertinent concernant le flou de mouvement. Une autre explication pourrait être l'âge moyen des observateurs. Sur les écrans 1 et 3, les observateurs ont un âge moyen de 25.5 ans (avec un écart-type de 7 ans), ce sont donc principalement de jeunes gens dans leur vingtaine. Les observateurs ayant participé aux tests sur l'écran 2 ont un âge moyen de 38 ans (avec un écart-type de 15 ans). L'âge peut-être un facteur important de la tolérance aux dégradations visuelles, la jeune génération, habituée aux vidéos bas-débit disponibles en streaming sur l'Internet, semble moins regardantes concernant la qualité générale d'un contenu vidéo. Enfin, il est intéressant de noter que le seuil de non acceptabilité pour les séquences « Texte » est relativement constant (moyenne de 9.2 minutes d'arc) pour les différents groupes d'observateurs. La lisibilité est apparemment une limite qui s'affranchit des différences culturelles et générationnelles.

Dans l'ensemble, la valeur moyenne du flou de mouvement perçu au seuil de non acceptabilité est de 11.7 minutes d'arc, néanmoins l'écart-type sur les 9 valeurs moyennes est de 2.53.

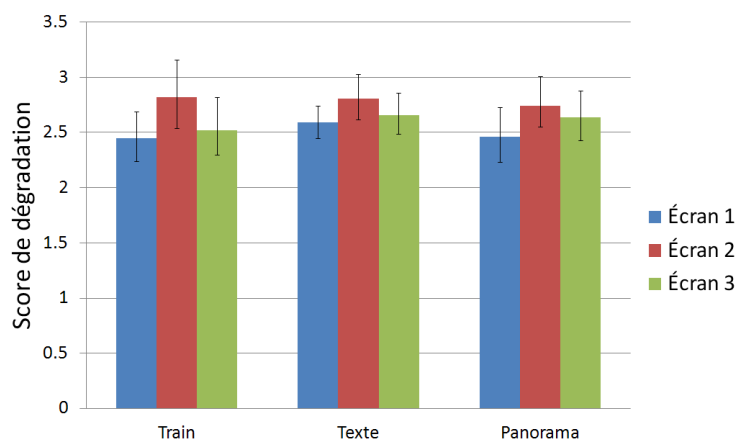


FIGURE 7.4 – Valeurs moyennes du score de gêne correspondant au seuil de non acceptabilité.

7.4.1.4 Score de gêne au seuil de non acceptabilité

La figure 7.4 présente les scores de gêne correspondant aux vitesses V_{accept} déterminées au seuil de non acceptabilité. Malgré les fortes différences observées sur le seuil de non acceptabilité, le score de gêne associé est relativement constant, pour les 3 écrans et pour les 3 types de séquences. Sa valeur moyenne est de 2.63, avec un faible écart-type sur les 9 valeurs moyennes de 0.14. Cette valeur de gêne se situe entre les qualificatifs « Légèrement gênant » et « Gênant » (cf. tableau 7.2).

Compte-tenu du protocole utilisé, il est normal que les observateurs aient tenté d'associer les deux échelles de notation. Il est intéressant de noter qu'en moyenne les observateurs considèrent comme « acceptable » un flou de mouvement « légèrement gênant ». Cette information est intéressante dans l'optique d'une évaluation de l'acceptabilité du flou de mouvement. Il semblerait qu'il ne soit pas nécessaire de supprimer totalement le flou de mouvement (chose par ailleurs très difficile à faire techniquement) pour contenter les utilisateurs d'écrans à cristaux liquides.

7.4.2 Conséquences sur la valeur recommandée du temps de flou perçu global

7.4.2.1 Caractéristiques statistiques du signal de télévision

Plusieurs études ont été faites récemment pour caractériser le signal vidéo diffusé en télévision [Fujine 07, van Heesch 08]. Ces travaux consistent à mesurer les caractéristiques statistiques du signal de télévision reçu par les utilisateurs sur de longues périodes de temps, et pour différents types de contenu.

Nous nous sommes intéressés ici aux vitesses moyennes observées dans les contenus télévisuels. Les valeurs rapportées par Funjine *et al.* [Fujine 07] et par Van Heesch *et al.* [van Heesch 08] sont similaires. Voici les principales constatations dont nous allons nous servir par la suite :

- Le flou de mouvement de caméra apparaît à partir de 10 deg/s.

- Les vitesses observées sont en majorité distribuées sous 20 deg/s, les vitesses maximales rencontrées sont de l'ordre de 30 deg/s.
- Les textes défilants (non sujet au flou de mouvement de caméra) ont une vitesse moyenne de 13.8 deg/s, avec quelques occurrences au delà de 20 deg/s.

7.4.2.2 Recommandations concernant le temps de flou global d'un écran LCD

À partir des valeurs que nous venons de présenter, nous pouvons maintenant déterminer des recommandations concernant le temps de flou perçu global d'un écran à cristaux liquides. Les prérequis sur le temps de flou perceptif global MB_{PERC} (cf. section 6.5) sont calculées de la façon suivante :

$$\begin{aligned} MB_{PERC} &\leq MB_{PERC_{lim}} \\ MB_{PERC_{lim}} &= \frac{BEW_{acc}}{V_{lim}} \end{aligned}$$

avec BEW_{acc} le niveau de flou acceptable pour le prérequis considéré, et V_{lim} la vitesse limite en deçà de laquelle nous considérons que l'apparition du flou de mouvement est inacceptable pour le prérequis considéré. Nous distinguons alors trois niveaux de prérequis concernant flou de mouvement sur les écrans à cristaux liquides, à chacun de ces prérequis correspond une valeur de vitesse limite V_{lim} et une valeur de flou acceptable BEW_{acc} .

Écran de qualité acceptable : L'écran sera considéré comme « acceptable » si de manière générale le flou perçu global est inférieur à la valeur moyenne de flou au seuil de non acceptabilité que nous avons mesurée précédemment ($BEW_{acc} = 11.7$ minutes d'arc), et ce pour des vitesses qui n'impliquent pas de flou de mouvement dû à la caméra ($V_{lim} = 10$ deg/s). Le temps de flou perceptif MB_{PERC} doit donc être inférieur à $MB_{PERC_{lim}} = 20$ ms.

Écran de bonne qualité : L'écran sera considéré de bonne qualité si, en plus de la condition précédente, le flou de mouvement perçu sur les textes défilants reste acceptable. Autrement dit, si pour une vitesse de défilement moyenne de 14 deg/s, la largeur de flou perçu n'excède pas 9.2 minutes d'arc ($V_{lim} = 14$ deg/s et $BEW_{acc} = 9.2$ arcmin). Le temps de flou perceptif MB_{PERC} doit donc être inférieur à $MB_{PERC_{lim}} = 11$ ms.

Écran d'excellente qualité : Enfin, l'écran sera considéré d'excellente qualité si, en plus des conditions précédentes, le flou de mouvement perçu global est inférieur au seuil de non acceptabilité ($BEW_{acc} = 11.7$ arcmin) pour la majorité des vitesses présentes dans le signal télévisuel, c'est à dire pour les vitesses inférieures à $V_{lim} = 30$ deg/s. Le temps de flou perceptif MB_{PERC} doit donc être inférieur à $MB_{PERC_{lim}} < 6.5$ ms.

Le tableau 7.9 récapitule les 3 niveaux de recommandation sur la valeur du temps de flou perceptif MB_{PERC} concernant le flou de mouvement sur les écrans à cristaux liquides. À titre d'exemple, les

	$MB_{PERC_{lim}}$
Écran de qualité acceptable	20 ms
Écran de bonne qualité	11 ms
Écran d'excellente qualité	6.5 ms

TABLE 7.9 – Recommandations concernant le temps de flou perceptif MB_{PERC} pour les écrans à cristaux liquides.

trois écrans utilisés pendant ces tests sont « acceptables », les écrans 2 et 3 sont assez proches du niveau « bon ». Cela correspond au ressenti des observateurs durant les tests.

7.5 Conclusion

Les travaux présentés dans ce chapitre ont consisté à évaluer l'acceptabilité du flou de mouvement, ainsi que la gêne associée à sa perception, pour différents types de séquences, et sur trois écrans. Les résultats de cette évaluation, combinés avec ceux des mesures psychophysiques de la perception du flou de mouvement décrites dans le chapitre 6, ont permis de déterminer des seuils en terme de largeur de flou de mouvement perçu. Trois seuils ont été déterminés : le seuil de perception du flou, le seuil de gêne et le seuil de non acceptabilité. Le seuil de non acceptabilité du flou de mouvement, c'est à dire la valeur de flou de mouvement perçu au delà de laquelle le flou devient inacceptable est supérieure au seuil de gêne. Autrement dit, il semble que les observateurs soient prêts à accepter une légère gêne visuelle liée au flou de mouvement.

La détermination de ces seuils a également permis d'émettre une recommandation concernant le valeur de flou de mouvement global représenté par le temps de flou perceptif MB_{PERC} défini dans le chapitre 6. En se basant, sur les statistiques des caractéristiques temporelles du signal vidéo diffusé en télévision, nous avons déterminé 3 valeurs seuil de temps de flou perceptif $MB_{PERC_{lim}}$: 20 ms, 11 ms et 6.5 ms, correspondant à un écran de qualité respectivement acceptable, bonne et excellente.

Quatrième partie

Impact du flou de mouvement sur la qualité visuelle

Chapitre 8

Évaluation subjective de la qualité visuelle sur les écrans LCD et CRT

Sommaire

8.1	Introduction	141
8.2	Tests d'évaluation subjective de la qualité	142
8.2.1	Méthodologie	142
8.2.2	Contenus évalués	143
8.2.3	Conditions d'observation	147
8.2.4	Observateurs	148
8.2.5	Résultats	149
8.3	Différence de qualité entre écran CRT et écran LCD	150
8.3.1	Cas des séquences vidéo	153
8.3.2	Cas des images fixes	155
8.3.3	Influence de la résolution	156
8.4	Conclusion	159

8.1 Introduction

Nous avons montré dans la partie précédente que le flou de mouvement introduit une sensation de gêne lorsqu'on visualise des séquences vidéo sur un écran à cristaux liquides. Nous allons à présent nous intéresser à l'influence de cette gêne sur la qualité visuelle de contenus naturels. Cette partie de nos travaux de thèse se base en grande partie sur les évaluations subjectives de la qualité visuelle menées durant la thèse de Stéphane Péchard au sein du laboratoire IRCCyN-IVC [Péchard 08], que nous avons étendues de manière à nous focaliser sur les différences de technologie d'affichage. En effet, nous nous intéressons ici plus particulièrement aux distorsions visuelles introduites par l'écran, et à la différence de qualité entre une séquence visualisée sur écran de type LCD, et sur écran de type CRT pour lequel le flou de mouvement est inexistant.

Dans ce chapitre, nous allons dans un premier temps décrire les tests d'évaluation subjective de la qualité qui ont été menés. Ceux-ci se sont déroulés à la fois sur écran CRT et sur écran LCD. Douze

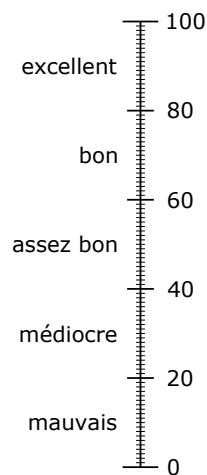


FIGURE 8.1 – Échelle continue d'évaluation de la qualité utilisée par la méthodologie SAMVIQ.

séquences vidéo et cinq images fixes non compressées, ainsi que leurs différentes versions dégradées, ont été évaluées durant cette campagne de tests. Les résultats de ces tests sont présentés en mettant l'accent sur les différences de qualité visuelle perçue sur les deux écrans.

8.2 Tests d'évaluation subjective de la qualité

8.2.1 Méthodologie

L'évaluation subjective de la qualité visuelle des séquences vidéo et des images fixes menée durant cette campagne de tests s'est déroulée en utilisant le protocole d'évaluation de la qualité visuelle SAMVIQ (*Subjective assessment methodology for video quality*). Ce protocole proposé par l'EBU [EBU 03] présente l'intérêt majeur de pouvoir discriminer des qualités proches les unes des autres [Blin 06].

SAMVIQ est une méthodologie d'évaluation à stimuli multiples et à échelle de qualité quasi continue. L'échelle de qualité est présentée figure 8.1. Elle comporte 100 notes différentes, allant de 1 à 100 par pas de 1. Les observateurs attribuent à chaque séquence une valeur numérique placée sur un segment de droite. Celui-ci est séparé en zones, étiquetées par des attributs sémantiques. Ces attributs permettent de donner une consistance à l'ensemble de l'échelle, afin d'éviter les différences d'utilisation.

Pour chaque contenu original, la méthodologie SAMVIQ permet une visualisation à accès aléatoire, c'est-à-dire que l'observateur choisit l'ordre dans lequel il visionne les versions traitées de ce contenu et peut suivre son propre rythme pour l'évaluation, la modification des notes et la répétition des séquences. Le nombre de visualisation pour chaque séquence/image n'est pas limité, c'est la dernière note attribuée qui est enregistrée. Par contre, chaque séquence/image doit être visualisée au moins

une fois entièrement avant de pouvoir être notée. Plusieurs contenus originaux, chacun ayant subi divers traitements sont évaluables par session. Lorsque l'observateur a terminé l'évaluation d'un contenu original et de l'ensemble des versions dégradées, il valide l'ensemble de ses évaluations et peut passer au contenu suivant.

Le principal avantage de cette méthodologie est de permettre à l'observateur d'affiner son jugement en regardant plusieurs fois les mêmes séquences/images, cela permet au final d'augmenter la précision de la mesure, et donc de diminuer le nombre d'observateurs nécessaire. SAMVIQ transforme le jugement de plusieurs observateurs ayant vu une fois un contenu en jugement de moins d'observateurs l'ayant vu plusieurs fois.

Pour chaque contenu, il est nécessaire de définir deux ancres sur l'échelle de qualité, l'un de haut niveau, l'autre de bas niveau. Cela permet d'exploiter l'échelle de qualité au maximum et de stabiliser la notation d'un observateur à l'autre. La séquence/image originale, non dégradée, est utilisée comme ancre de haut niveau : c'est la référence explicite. L'ancre de bas niveau est défini par une version codée à bas débit. Pour chaque contenu à évaluer, l'observateur a donc accès simultanément à une série de séquences/images qu'il peut afficher dans l'ordre qu'il souhaite et autant de fois qu'il le veut. Cette série est constituée des différentes versions traitées que l'on cherche à évaluer, parmi lesquelles se trouvent la version d'ancre bas niveau et une version cachée de la référence non traitée. La référence explicite est elle clairement désignée comme la version de référence du contenu en cours d'évaluation. Au final, pour un contenu donné l'observateur doit évaluer entre 5 et 12 versions traitées (référence cachée, et ancre bas compris). Typiquement, quatre contenus sont évalués durant une session.

8.2.2 Contenus évalués

Différents types de contenus ont été utilisées durant cette campagne d'évaluation subjective. L'objectif principal de cette étude et de mesurer l'influence du flou de mouvement sur LCD sur la qualité visuelle d'une séquence vidéo. Pour cette raison, nous avons mené les évaluations subjectives de la qualité sur des séquences avec différentes quantité de mouvement mais aussi sur des images fixes.

Différentes versions des contenus sélectionnés ont été présentées aux observateurs : les versions non dégradées (non compressées) et plusieurs versions dégradées (codées).

8.2.2.1 Séquences vidéo

Douze séquences vidéo ont été utilisées durant cette campagne de tests. Toutes ces séquences sont au format haute-définition entrelacé 1080i, cadencé à 50 trames par seconde. Elles sont par conséquent destinées à être affichées sur un écran haute-définition d'une résolution de 1920×1080 pixels.

Ces séquences sont réparties en trois ensembles, définis par la provenance et le format des séquences.

Groupe « SVT » Le premier groupe est constitué de quatre séquences vidéo réalisées par la chaîne de télévision suédoise SVT en 2002 [SVT 02]. Ces vidéos sont librement mises à la disposition de



(a) *New Mobile and Calendar* : en intérieur, mouvements multiples.



(b) *Parkrun* : en extérieur, traveling horizontal rapide.



(c) *Knightshields* : en intérieur, traveling horizontal puis zoom.



(d) *Stockholm Pan* : en extérieur, panorama, mouvement moyen.

FIGURE 8.2 – Vidéos HDTV du groupe « SVT ».

la communauté, et ont spécifiquement été tournées pour la recherche scientifique. Les valeurs de chaque pixel sont codées dans l'espace colorimétrique YC_bC_r , avec un échantillonnage couleur 4:2:0. Ces contenus ont une durée de 10 secondes, soit 500 trames. La figure 8.2 présente, pour chaque contenu, une image extraite de la séquence vidéo, ainsi qu'une courte description.

Groupes « Euro1080-1 » et « Euro1080-2 » Les deux autres groupes sont constitués chacun de quatre séquences réalisées par la chaîne de télévision belge Euro1080. Ces contenus ne sont pas libres de droit, ils ne peuvent donc être utilisés que dans un but scientifique, sans pouvoir être redistribués. Les valeurs de chaque pixel sont codées dans l'espace colorimétrique YC_bC_r , avec un échantillonnage couleur 4:2:2.

La durée des séquences de ce sous-ensemble est variable. Quatre séquences ont une durée de 12 secondes (soit 600 trames), elles constituent le groupe Euro1080-1. Les quatre autres séquences ont une durée de 9 secondes (soit 450 trames), c'est le groupe Euro1080-2. La figure 8.3 présente, pour chaque contenu, une image extraite de la séquence vidéo, ainsi qu'une courte description.

Génération des séquences dégradées Pour chacune des douze séquences, huit versions différentes ont été évaluées par les observateurs : la version non codée (et donc non dégradée) et sept versions compressées (et donc dégradées).

Les séquences dégradées ont été générées par le codeur de référence H.264 [JVT 05]. Le profil principal (*main profile*) et le niveau 4.0 (*level 4.0*) ont été utilisés pour coder les quatre séquences SVT. Les séquences Euro1080 ont été codées en utilisant le profil étendu (*high profile*), car le profil principal ne gère pas le 4:2:2, et le niveau 4.0.

Les sept débits utilisés pour chaque séquence ont été définis par un panel d'experts. La gamme de débits est différente d'un contenu à un autre. Ceci est dû aux différences de complexité spatio-temporelle de chaque contenu. Les valeurs de débit ont été choisies de manière à ce que la qualité des séquences générées couvre l'ensemble de l'échelle catégorielle de qualité, c'est à dire de « mauvais » à « excellent ». Les débits choisis, ainsi que l'ensemble des informations utiles concernant les douze séquences, sont récapitulés dans le tableau 8.1.



(a) *Concert* (12 secondes) : en intérieur, faible luminance, mouvements moyens.



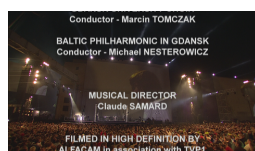
(b) *Foot* (12 secondes) : en extérieur, traveling horizontal puis plan sur les joueurs.



(c) *Movie* (12 secondes) : en intérieur, traveling avant puis zoom.



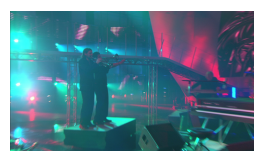
(d) *Voile* (12 secondes) : en extérieur, plusieurs plans.



(e) *Credits* (9 secondes) : en intérieur, faible luminance, défilement de texte, plusieurs changements de plan derrière le texte.



(f) *Golf* (9 secondes) : en extérieur, peu de mouvement.



(g) *Show* (9 secondes) : en intérieur, zoom puis plan fixe, peu de mouvement.



(h) *Standing* (9 secondes) : en intérieur, plan fixe, très peu de mouvement.

FIGURE 8.3 – Vidéos HDTV des groupes « Euro1080-1 » (première ligne) et « Euro1080-2 » (seconde ligne).

Lors des tests d'évaluation, les observateurs ont donc évalué 9 séquences pour chacun des 12 contenus vidéo : la référence explicite (non traitée) et 8 versions non étiquetées constituées des 7 versions codées et de la référence cachée (non traitée également). La version d'ancrage de niveau bas correspond à la version codée avec le débit le plus bas. Au total, ce sont donc 96 séquences qui ont été évaluées durant cette campagne de tests, sur chacun des deux types d'écran. Chaque groupe est constitué de 32 séquences.

8.2.2.2 Images fixes

De manière à mesurer également la différence de qualité visuelle entre les deux types d'écrans en l'absence de mouvement, des images fixes ont également été évaluées. La figure 8.4 présente les cinq images utilisées. Ces images, non compressées, ont une résolution de 1280×800 pixels. Ce sont des images naturelles représentant différents types de contenus.

Comme pour les séquences vidéo, ces cinq images ont été traitées de manière à obtenir plusieurs versions d'une même image et pour forcer les observateurs à construire leur jugement sur l'ensemble de l'échelle de qualité. En plus de la version originale non dégradée (référence explicite), les observateurs ont évalué 4 autres versions de la même image : la référence cachée, une version d'ancrage bas niveau obtenue par un codage JPEG avec un taux de compression élevé, deux versions obtenues par un redimensionnement d'une version basse résolution de l'image (640×400 pixels), respectivement avec une interpolation au plus proche voisin et une interpolation bilinéaire.

Au total, le groupe « Images » est donc constitué des 25 images fixes.

Nom	Durée	Échantillonnage	Débits des versions dégradées (Mbps)
<i>New Mobile & Calendar</i>	10 s	4 :2 :0	2.25 - 2.5 - 3.15 - 4 - 5 - 7 - 10
<i>Parkrun</i>	10 s	4 :2 :0	8 - 12 - 16 - 18 - 20 - 24 - 28
<i>Knightshields</i>	10 s	4 :2 :0	2.25 - 3 - 4 - 5 - 6 - 7 - 8
<i>Stockholm Pan</i>	10 s	4 :2 :0	1.625 - 1.875 - 2.25 - 3 - 3.6 - 4 - 6
<i>Concert</i>	12 s	4 :2 :2	4 - 8 - 10 - 14 - 18 - 24 - 30
<i>Foot</i>	12 s	4 :2 :2	4 - 6 - 10 - 14 - 18 - 22 - 26
<i>Movie</i>	12 s	4 :2 :2	2.5 - 4 - 7 - 8 - 9 - 10 - 14
<i>Voile</i>	12 s	4 :2 :2	2.5 - 4 - 6 - 8 - 9 - 11 - 12
<i>Credits</i>	9 s	4 :2 :2	4 - 6 - 7 - 8 - 9 - 10 - 14
<i>Golf</i>	9 s	4 :2 :2	1 - 1.5 - 2 - 2.5 - 3 - 4 - 6
<i>Show</i>	9 s	4 :2 :2	2 - 3 - 4 - 8 - 10 - 14 - 22
<i>Standing</i>	9 s	4 :2 :2	0.75 - 1.25 - 2 - 3 - 5 - 6 - 8

TABLE 8.1 – Tableau récapitulatif des caractéristiques des douze séquences utilisées.

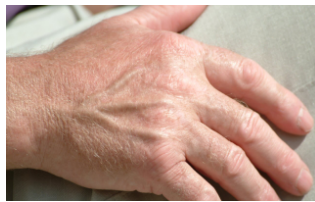
(a) *Football.*(b) *Hand.*(c) *House.*(d) *Landscape.*(e) *Map.*

FIGURE 8.4 – Groupe « Images ».

Type	Hauteur de l'image affichée	Formats supportés	Luminance maximale
CRT	20.5 cm	1080i, 720p	71 cd/m ²
LCD	46 cm	1080p	440 cd/m ²

TABLE 8.2 – Caractéristiques des deux écrans utilisés.

8.2.3 Conditions d'observation

8.2.3.1 Écrans

L'évaluation subjective de la qualité visuelle des séquences vidéo et des images fixes a été réalisée sur deux écrans. Tous deux sont destinés à afficher des contenus TV haute-définition.

Le premier est un moniteur professionnel de type CRT, il peut afficher deux types de format haute-définition : le 1080i (1920 × 1080 en mode entrelacé) et le 720p (1280 × 720 en mode progressif).

Le second est un écran de télévision de type LCD adapté à nos besoins. En effet, les traitements permettant l'adaptation de différents types de contenus d'entrée au format du panneau de pixels proprement dit ont été désactivés, ainsi que les divers procédés d'amélioration de l'image qui existent sur la plupart des téléviseurs LCD. De cette manière, nous connaissons exactement la nature de l'image affichée par l'écran, vierge des éventuelles distorsions apportées par tous ces traitements. Une des conséquences de ce choix technologique est que cet écran ne peut supporter qu'un seul format d'entrée, le 1080p (1920 × 1080 en mode progressif). Il convient également de noter que cet écran LCD correspond à l'écran n°3 utilisé durant les expérimentations présentées dans les chapitre 6 et 7.

Le tableau 8.2 présente les principales caractéristiques des deux écrans utilisées pour ces tests.

8.2.3.2 Salle de tests

Les tests se sont déroulés dans une salle spécifique, dont les murs sont recouverts d'un tissu neutre. Les conditions de lumière et la distance d'observation ont été mesurés et ajustés suivant les recommandations de l'ITU [ITU 98, ITU 04]. Pour des contenus haute-définition la distance d'observation recommandée est égale à trois fois la hauteur de l'image. Par conséquent, la distance a été fixée à 61.5 cm pour l'écran CRT, et à 138 cm pour l'écran LCD.

8.2.3.3 Adaptation des différents types de contenus aux deux écrans

Désentrelacement des séquences sur l'écran LCD Les 12 séquences utilisées sont au format 1080i. Leur visualisation sur l'écran CRT ne pose aucun problème puisque celui-ci accepte ce type de format en entrée. Sur l'écran LCD cependant, seul un format progressif est accepté.

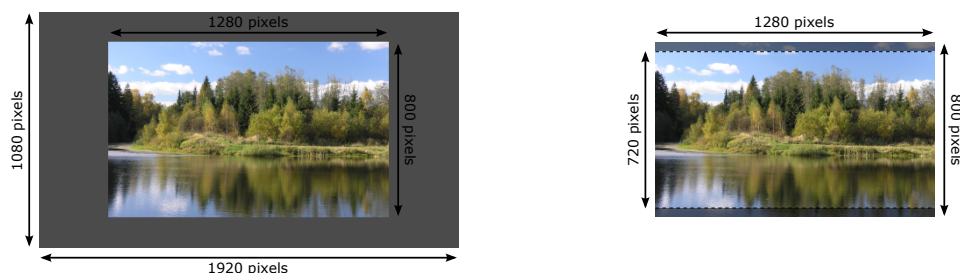
Pour pouvoir être affichés sur l'écran LCD, les contenus vidéo ont donc été désentrelacés. Le procédé de désentrelacement consiste à passer d'un format entrelacé à un format progressif en interpolant les lignes paires (ou impaires) manquant dans chaque trame. Cette opération est difficile à réaliser, elle est généralement effectuée par l'intermédiaire de traitements internes à l'écran. Ces traitements ayant été volontairement désactivés sur notre écran LCD, nous avons utilisé une unité externe de désentrelacement qui permet d'effectuer le désentrelacement du signal d'entrée en temps réel, en amont de l'écran. La qualité du désentrelacement a été testée par plusieurs experts en comparant une séquence vidéo progressive avec la même séquence entrelacée puis désentrelacée. Aucune différence n'a été observée. Il convient néanmoins de rester prudent car certains types de contenus peuvent mettre le désentrelaceur en défaut, et des distorsions visuelles peuvent apparaître sur le signal affiché.

Format des images fixes Les images fixes utilisées ont une résolution de 1280×800 pixels. Deux problèmes se posent alors. Premièrement, l'écran LCD n'accepte qu'un format d'entrée de 1920×1080 pixels. Deuxièmement, pour pouvoir comparer les résultats de l'évaluation des images et de l'évaluation des séquences vidéo, les conditions de tests doivent être les mêmes, notamment la distance d'observation (ou plus précisément le rapport pixels par degré visuel, qui doit rester égal à 57 pixels/deg). Une solution aurait pu être de redimensionner les images mais cela aurait introduit des distorsions visuelles, ce qui n'était pas souhaitable. Nous avons par conséquent décidé d'augmenter la taille des images, non pas en les redimensionnant mais en ajoutant une bordure grise sur les quatre côtés de l'image originale, l'image finale possède une taille de 1920×1080 pixels. Cette opération est illustrée figure 8.5a. La distance d'observation reste la même que pour l'évaluation des séquences vidéo (138 cm). Cela correspond pour ces images à une distance d'environ 4 fois la hauteur de l'image (image originale sans bordure).

Sur l'écran CRT, les mêmes images sont utilisées. Ici encore la distance de visualisation reste la même que pour l'évaluation des vidéos (61.5 cm), soit 4 fois la hauteur de l'image sans la bordure grise. À cette résolution spatiale, les images ne peuvent être affichées que dans un mode entrelacé (cf. tableau 8.2). Lors de l'affichage d'une image fixe en mode entrelacé, le clignotement introduit par le balayage de l'écran CRT peut être visible par l'observateur. Par conséquent, une autre solution a également été testée (seulement sur l'écran CRT) : l'affichage au format 720p. Pour ce faire, deux bandes horizontales de 40 pixels ont été rognées, en haut et en bas, sur l'image originale (figure 8.5b). L'image obtenue possède alors une résolution de 1280×720 pixels et peut être affichée en mode progressif. Dans ce cas là, la distance d'observation est ajustée à 4.5 fois la hauteur de l'image (égale à la hauteur de l'écran), soit 92.3 cm, de manière à garder une résolution visuelle constante pour l'ensemble des contenus testés (57 pixels par degré visuel).

8.2.4 Observateurs

L'ensemble des évaluations a été divisé en 9 sessions de tests, chaque session correspond à l'évaluation d'un groupe de contenus, sur un type d'écran donné. Le tableau 8.3 décrit les 9 sessions de tests.



(a) Image originale insérée dans une image grise de 1920×1080 pixels, pour être affichée au format 1080p sur LCD et au format 1080i sur CRT. (b) Image originale rognée au format 1280×720 pixels, pour être affichée au format 720p sur CRT.

FIGURE 8.5 – Adaptation du format des images fixes au format des écrans de tests.

Session	Écran	Format	Groupe de contenus	Nb contenus	Nb obs (retenus/total)
1	CRT	1080i	SVT	36	28/32
2	LCD	1080p	SVT	36	22/24
3	CRT	1080i	Euro1080 - 1	36	29/34
4	LCD	1080p	Euro1080 - 1	36	28/29
5	CRT	1080i	Euro1080 - 2	36	28/32
6	LCD	1080p	Euro1080 - 2	36	34/39
7	CRT	1080i	Images	20	15/18
8	CRT	720p	Images	20	17/18
9	LCD	1080p	Images	20	17/18

TABLE 8.3 – Les 9 sessions d'évaluation subjective de la qualité visuelle.

À chaque session correspond un panel différent d'observateurs. Chacun de ces observateurs a subi un examen de manière à vérifier son acuité visuelle par le test de Monoyer [Monoyer 75] et l'absence de daltonisme par les tests d'Ishihara [Ishihara 67].

Pour chaque session de tests, la cohérence des résultats de chaque observateur est évaluée, de manière à détecter, et à rejeter le cas échéant, les observateurs non consistants. La méthodologie SAMVIQ propose un rejet des observateurs basé sur une comparaison des notes de qualité données par un observateur avec les notes moyennes sur l'ensemble des observateurs. Deux critères sont calculés, le coefficient de corrélation linéaire et le coefficient de corrélation de rang. Si l'un des deux coefficients est inférieur à un certain seuil, l'observateur est rejeté. Nous avons fixé le seuil à 0.85. Le tableau 8.3 donne le nombre d'observateurs retenus après rejet, pour chaque session. Le nombre minimum d'observateurs recommandé par la méthodologie SAMVIQ est de 15 [EBU 03].

8.2.5 Résultats

Les résultats de ces évaluations subjectives de la qualité sont présentées dans les figures 8.6 à 8.8. Sur chaque graphe, la note de qualité moyenne Q sur l'ensemble des observateurs est tracée en fonction du débit de la séquence codée, ou du traitement dégradant réalisé sur les images fixes. La

note moyenne obtenue par la référence cachée est également donnée (abscisse REF). Les intervalles de confiance à 95% sont donnés par les segments verticaux associés à chaque point.

Pour chaque contenu, les résultats sont donnés à la fois pour l'écran CRT et pour l'écran LCD. Pour les images fixes (figure 8.7), deux séries de résultats ont été obtenues pour l'écran CRT : la série « CRT (1080i) » correspond aux images insérées dans une image grise de 1920×1080 pixels (cf. figure 8.5a), tandis que la série « CRT (720p) » correspond aux images rognées (cf. figure 8.5b). Un décalage horizontal a été inséré entre chaque série pour faciliter la lecture des courbes.

8.3 Différence de qualité entre écran CRT et écran LCD

On observe sur les figures 8.6 à 8.8 que des différences significatives existent entre les notes de qualité obtenues sur écran CRT, Q_{CRT} et les notes de qualité obtenues sur écran LCD, Q_{LCD} . On note ΔQ la différence entre la note de qualité moyenne obtenue sur CRT et la note de qualité moyenne obtenue sur LCD, pour chaque contenu i et pour chaque traitement j (codage H.264 à différents débits sur les séquences, codage JPEG et redimensionnement sur les images fixes, contenus de référence non dégradés dans les deux cas) :

$$\Delta Q_{i,j} = Q_{\text{CRT},i,j} - Q_{\text{LCD},i,j} \quad (8.1)$$

On s'intéresse ici à la perte de qualité due à l'affichage sur écran à cristaux liquides (par rapport à la qualité obtenue sur un écran CRT, qui peut être considérée comme la qualité de référence). Pour cela, nous distinguons la perte de qualité mesurée sur les séquences de référence, non dégradées. Celle-ci est notée ΔQ_{REF} , on obtient une valeur de ΔQ_{REF} par contenu i :

$$\Delta Q_{\text{REF}} = Q_{\text{CRT},i,j=\text{REF}} - Q_{\text{LCD},i,j=\text{REF}} \quad (8.2)$$

Nous avons également choisi de nous intéresser, pour chaque contenu i , à la différence de qualité ΔQ pour les N_i traitements j ayant obtenu une note de qualité supérieure ou égale à 70 sur l'écran CRT. Cette note de qualité correspond au centre de la catégorie « bon ». On peut donc considérer que dans cette région de l'échelle de qualité les distorsions liées au codage restent négligeable vis-à-vis des distorsions dues à l'écran. La moyenne des différences de qualité entre CRT et LCD pour l'ensemble des traitements j vérifiant cette condition est notée ΔQ_{MOY} , on obtient une valeur de ΔQ_{MOY} par contenu i :

$$\Delta Q_{\text{MOY}} = \frac{1}{N_i} \sum_{j \mid Q_{\text{CRT},i,j} \geq 70} (Q_{\text{CRT},i,j} - Q_{\text{LCD},i,j}) \quad (8.3)$$

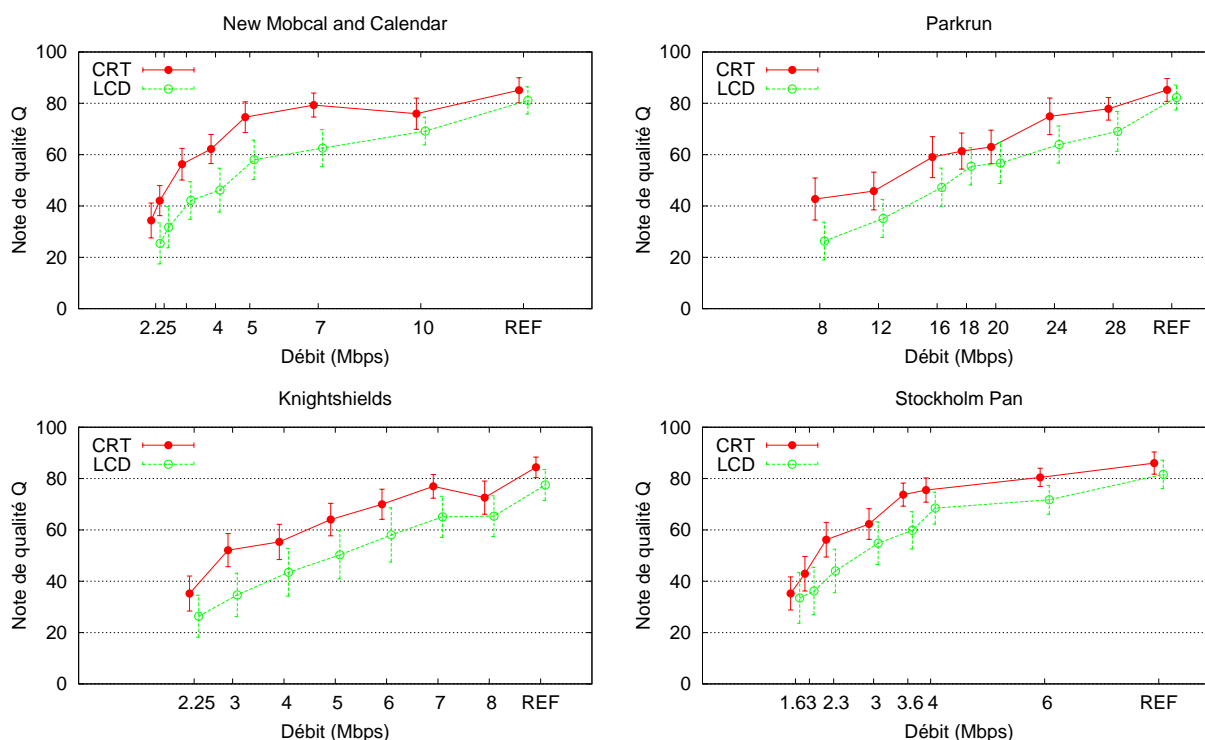


FIGURE 8.6 – Résultats de l'évaluation subjective des séquences SVT, pour chaque débit et sur chaque type d'écran.

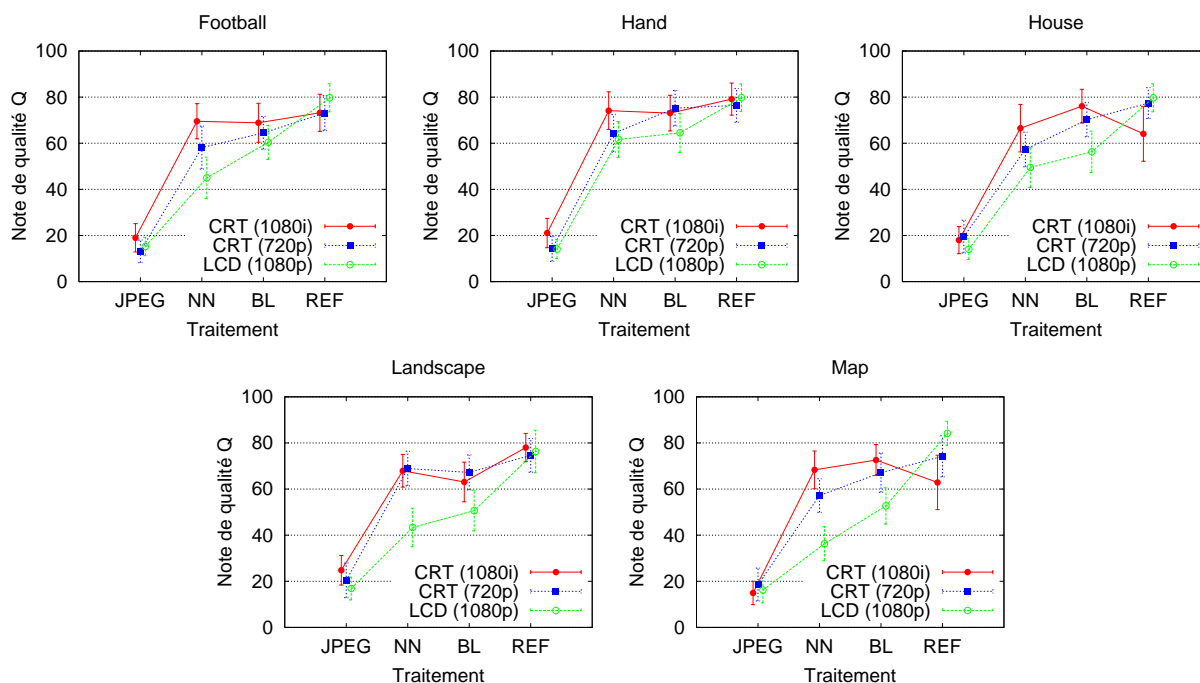


FIGURE 8.7 – Résultats de l'évaluation subjective des images fixes, pour chaque traitement et sur chaque type d'écran. Les deux formats 1080i et 720p ont été testés sur l'écran CRT. « NN » correspond à un redimensionnement par interpolation au plus proche voisin, « BL » par interpolation bilinéaire.

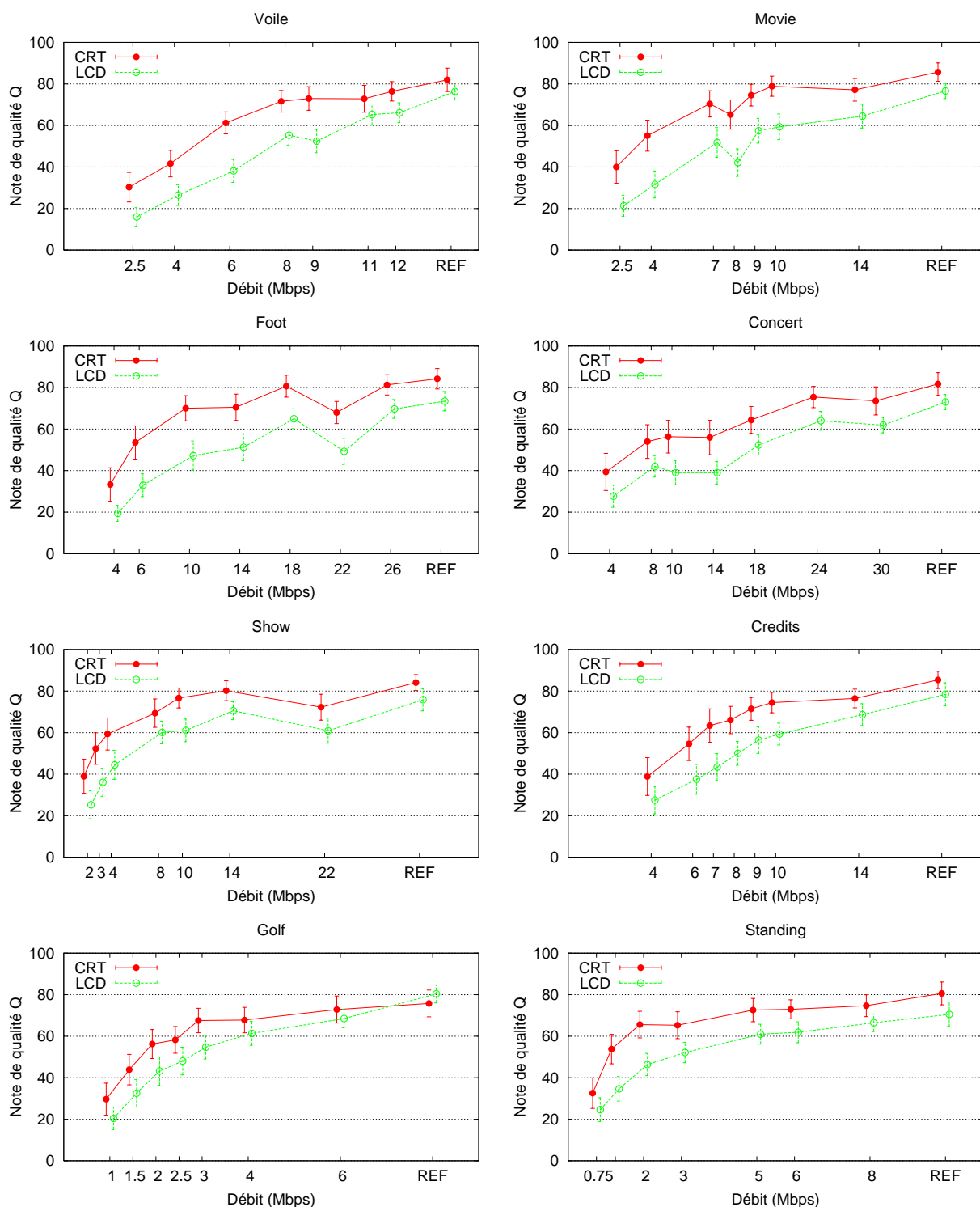


FIGURE 8.8 – Résultats de l'évaluation subjective des séquences Euro1080, pour chaque débit et sur chaque type d'écran.

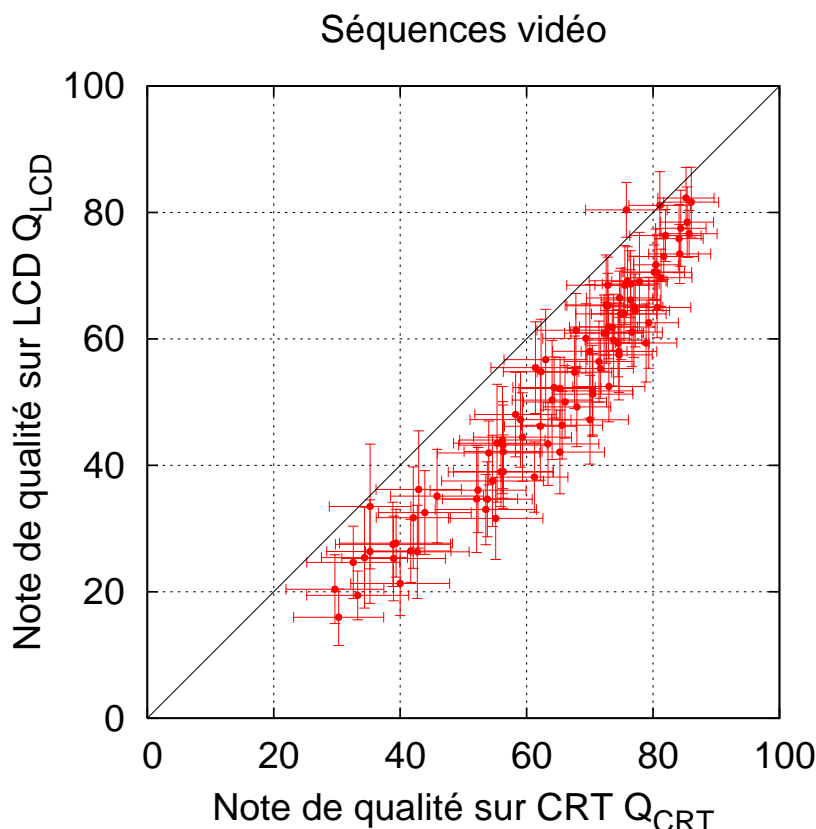


FIGURE 8.9 – Notes de qualité moyennes obtenues sur LCD Q_{LCD} en fonction des notes de qualité moyennes obtenues sur CRT Q_{CRT} , pour les 96 séquences vidéo testées (12 références et leurs 7 versions dégradées).

8.3.1 Cas des séquences vidéo

La figure 8.9 présentent les notes de qualité moyenne obtenues sur LCD Q_{LCD} en fonction des notes de qualité moyennes obtenues sur CRT Q_{CRT} , pour les 96 séquences testées (12 séquences vidéo de références et leurs 7 versions dégradées respectives). Tout d'abord, on peut noter qu'il existe une forte corrélation entre les notes obtenues sur LCD et celles obtenues sur CRT. Le coefficient de corrélation linéaire entre les deux séries de notes est égale à 0.9521.

$\Delta Q_{i,j}$ est toujours positif, excepté pour 2 séquences (sur un total de 96). La valeur moyenne de cette perte de qualité sur l'ensemble des séquences est égale à 13.3, soit plus de 10% de l'échelle de qualité. À quelques rares exceptions près, cette différence de qualité est statistiquement significative en regard des intervalles de confiance à 95%.

ΔQ_{REF} et ΔQ_{MOY} sont données pour chaque séquence dans le tableau 8.4. Les valeurs moyennes de ces deux mesures sur les 12 contenus sont également données. La perte de qualité moyenne sur les séquences ayant obtenue un score de qualité supérieur à 70 sur CRT a une valeur moyenne de 10.2, tandis que la perte de qualité sur les séquences de référence a une valeur moyenne de

Séquence	ΔQ_{MOY}	ΔQ_{REF}
<i>New Mobcal</i>	10.0	4.0
<i>Parkrun</i>	7.6	2.9
<i>Knightshields</i>	9.5	6.9
<i>Stockholm Pan</i>	8.5	4.4
<i>Voile</i>	12.1	5.6
<i>Movie</i>	15.4	9.0
<i>Foot</i>	16.0	10.8
<i>Concert</i>	10.6	8.7
<i>Show</i>	11.2	8.3
<i>Credits</i>	11.2	6.9
<i>Golf</i>	-0.1	-4.6
<i>Standing</i>	10.3	10.1
Moyenne	10.2	6.1

TABLE 8.4 – Différence de qualité ΔQ entre CRT et LCD pour chaque séquence, ΔQ_{MOY} est la moyenne sur les séquences ayant obtenues une note moyenne de qualité supérieure à 70 sur l'écran CRT, ΔQ_{REF} est la différence de qualité pour la seule séquence de référence.

6.1. Néanmoins, on peut observer une très grande variabilité d'un contenu à l'autre. Les valeurs extrêmes étant obtenue pour la séquence *Golf* d'une part avec $\Delta Q_{\text{REF}} = -4.6$ et $\Delta Q_{\text{MOY}} = -0.1$ (séquence globalement préférée sur LCD) et pour la séquence *Foot* d'autre part avec $\Delta Q_{\text{REF}} = 10.8$ et $\Delta Q_{\text{MOY}} = 16.0$ (séquence globalement très préférée sur CRT).

La perte de qualité sur LCD semble être liée à la quantité et à l'amplitude des mouvements présents dans les séquences. Pour les séquences *Foot* et *Movie*, ainsi que *Concert*, *Show*, *Credits* et *Voile* dans une moindre mesure, la différence de qualité est assez élevée, or ce sont des séquences présentant de forts mouvements. À l'inverse, la séquence *Golf* présente très peu de mouvement, or pour cette séquence les deux mesures de la perte de qualité sur LCD sont négatifs, indiquant que pour cette séquence les notes de qualité obtenues sur LCD ont été supérieures à celles obtenues sur CRT.

Ces premières observations montrent une certaine cohérence entre mouvements et perte de qualité sur l'écran LCD. Il semble donc que la différence de qualité observée entre les deux écrans puissent être due à la présence de mouvement dans la séquence. Il existe cependant des exceptions. La perte de qualité sur LCD est élevée pour la séquence *Standing*, alors que cette séquences ne présente que peu de mouvements. Cette différence peut alors s'expliquer par d'autres facteurs. La séquence *Standing* est constituée d'un arrière-plan très sombre, sur lequel le bruit d'acquisition est assez visible, particulièrement sur l'écran LCD dont la dynamique en luminance est environ 6 fois plus élevée.

En conclusion, on observe une différence significative entre la qualité visuelle mesurée sur CRT et celle mesurée sur LCD. Cette perte de qualité sur LCD s'observe aussi bien pour des contenus dégradés par un codage H.264 que pour des contenus non compressés. Pour ces derniers, seules les distorsions liées à l'affichage peuvent expliquer la différence de qualité observée. La quantité de mouvement présent dans chaque séquence semble être corrélée avec la perte de qualité ΔQ .

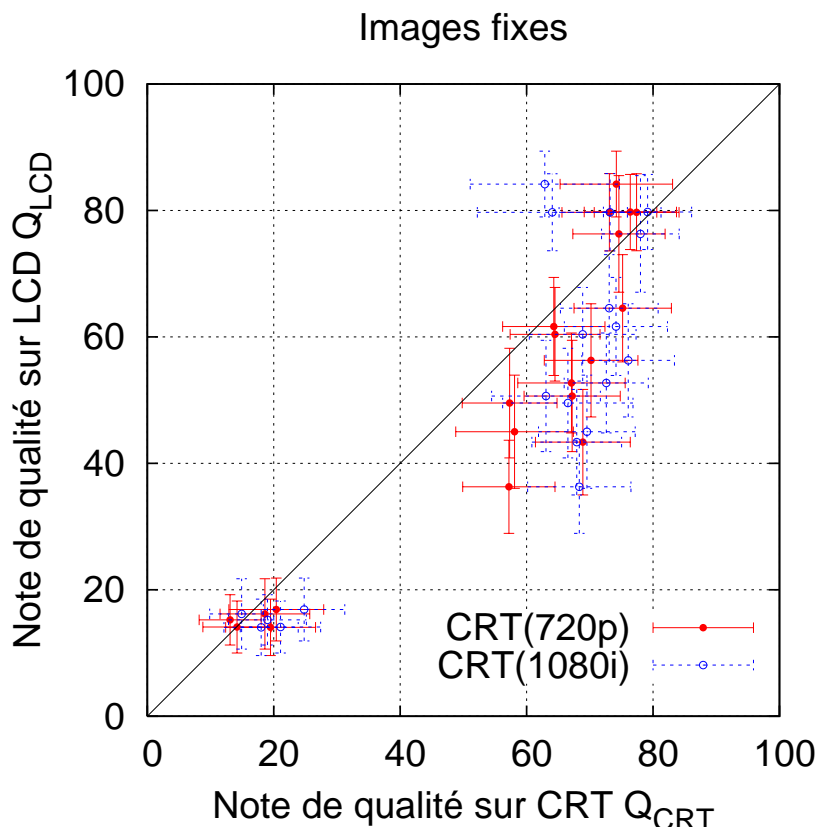


FIGURE 8.10 – Notes de qualité moyennes obtenues sur LCD Q_{LCD} en fonction des notes de qualité moyennes obtenues sur CRT Q_{CRT} , pour les 20 images fixes testées (5 références et leurs 3 versions dégradées respectives).

8.3.2 Cas des images fixes

La figure 8.10 présente les notes de qualité moyennes obtenues sur LCD Q_{LCD} en fonction des notes de qualité moyennes obtenues sur CRT Q_{CRT} pour les 20 images fixes testées (5 images de référence et leur 3 versions dégradées respectives). Deux séries de points sont représentées, la première correspond aux Q_{CRT} obtenues avec un affichage au format 720p, la seconde correspond au format d’affichage 1080i.

Sur cette figure, on distingue nettement trois nuages de points, qui correspondent respectivement aux images de référence non dégradées (qualité bonne à excellente), aux images dégradées par redimensionnement (qualité médiocre à bonne), et enfin aux images d’ancrage basse dégradées par codage JPEG (qualité mauvaise). Les différences entre la série de notes obtenues sur LCD et les deux séries de notes obtenues sur CRT varient selon le type de dégradation considérée.

Les images dégradées par une forte compression JPEG ont servi d’ancrage basse pour le jugement des observateurs, pour les inciter à utiliser l’ensemble de l’échelle de qualité. On remarque que les notes obtenues sont statistiquement similaires entre les deux écrans, avec une moyenne globale inférieure

Image	ΔQ_{REF} (1080i)	ΔQ_{REF} (720p)
<i>Football</i>	-6.5	-6.6
<i>Hand</i>	-0.6	-3.4
<i>House</i>	-15.6	-2.4
<i>Landscape</i>	1.7	-1.6
<i>Map</i>	-21.3	-10.0
Moyenne	-8.5	-4.8

TABLE 8.5 – Différence de qualité ΔQ_{REF} pour les versions de référence de chaque images fixes testées. Les valeurs sont données pour les deux formats d’affichage sur CRT testés (1080i et 720p).

à 20.

Concernant les images dégradées par redimensionnement, les notes de qualité obtenues sur CRT sont significativement meilleures. De plus les notes obtenues sur CRT avec le format d’affichage entrelacé 1080i sont en général légèrement meilleures que celles obtenues sur le même écran avec le format d’affichage progressif. La forte préférence pour l’écran CRT dans le cas de ces deux traitements par redimensionnement peut s’explique par le fait que le pixel d’un écran CRT est légèrement étalé spatialement tandis que le pixel d’un LCD est plus « net ». Par conséquent, les distorsions sont atténuées sur les écrans CRT alors qu’elles ont tendance à être plus visibles sur les écrans LCD.

Enfin, les cinq images de référence ont toutes été préférées sur l’écran LCD, excepté l’image *Landscape* qui a été préférée sur CRT au format 1080i. Les notes obtenues sur CRT sont équivalentes en mode progressif et en mode entrelacé, sauf pour les contenus *House* et *Map* qui contiennent beaucoup de contours horizontaux sur lesquels le clignotement du balayage est plus visible. Pour ces deux images, les notes obtenues avec le format d’affichage 1080i sont très faibles. Le tableau 8.5 donne la valeur de ΔQ_{REF} pour chacun des cinq contenus, et pour les deux formats d’affichage sur CRT. La différence de qualité moyenne pour un affichage au format 1080i est de -8.5, elle est de -4.8 au format 720p.

En conclusion, on observe que les notes de qualité obtenues sur l’écran LCD sont meilleures que celles obtenues sur l’écran CRT lorsqu’on considère les images de référence, qui ne possèdent d’autres distorsions que celles apportées par l’affichage. Il n’y a donc pas de perte de qualité sur LCD pour les images fixes. Ce résultat va dans le même sens que les résultats obtenus pour les séquences vidéo : le mouvement semble être responsable de la perte de qualité sur écran LCD. On peut également noter que les différences de qualité entre le format d’affichage 1080i et le format 720p sur CRT dépendent fortement de l’image considérée. Il y a très peu de différences pour les images *Football*, *Hand* et *Landscape*, tandis que la différence est importante pour *House* et *Map* car leurs nombreux contours horizontaux en font des images mal adaptées à l’affichage en mode entrelacé.

8.3.3 Influence de la résolution

Les résultats obtenus aussi bien sur les séquences vidéo que sur les images fixes laissent présager que le mouvement joue un facteur important de la perte de qualité observée sur les écrans à cristaux

liquides. De manière à confirmer cette hypothèse, nous avons décidé de réaliser de nouveaux tests d'évaluation subjective de la qualité, avec les mêmes séquences, non plus au format TVHD mais au format TVSD. En effet, pour une séquence donnée, l'amplitude des mouvements présents dans la séquence est fonction de la résolution. Si la résolution est divisée par deux, alors l'amplitude des mouvements mesurée en pixels sera également divisée par deux. Or, conformément aux recommandations, la distance de visualisation est elle aussi fonction de la résolution, ce qui a pour conséquence de maintenir constant le nombre de pixels affichés dans un degré visuel. Par conséquent l'amplitude des mouvements est également divisée par deux dans le champ visuel. Si comme nous le pensons le mouvement est bien le facteur déterminant de la perte de qualité observée sur les écrans LCD, alors des tests sur des séquences de résolution standard confirmeront ce résultat.

Pour ces tests, nous avons utilisé les quatre séquences de la base SVT. Ces séquences au format TVHD 1080i ont été transformées par un filtrage passe-bas demi-bande suivi d'un sous-échantillonnage par deux dans les deux directions. Le format obtenu est donc un format entrelacé de 960×540 pixels. Ce format 540i est assez proche du format TVSD (720×576 pixels en mode entrelacé), nous l'appelons ici QHD pour « quart de HD ». Nous avons préféré le format QHD au format TVSD car ce dernier nécessitait une interpolation des valeurs de pixels, ce qui aurait invariablement introduit de nouvelles dégradations visuelles.

À partir de ces quatre séquences originales au format QHD, 6 séquences dégradées ont été générées grâce au codeur H.264 dans les mêmes conditions que celles décrites dans la section 8.2.2.1. Ces 28 séquences ont été évaluées par les observateurs sur les deux types d'écran, avec la méthodologie SAMVIQ. Les conditions de visualisation sont les mêmes que celles décrites précédemment, exceptée la distance d'observation qui est doublée, passant de 3H pour les séquences HD à 6H dans le cas des séquences QHD (conformément aux recommandations [ITU 04]). 25 observateurs ont participé aux tests sur l'écran CRT et 18 sur l'écran LCD. Après rejet des observateurs incohérents, respectivement 22 et 16 observateurs ont été retenus. Les notes de qualité moyennes sur l'ensemble des observateurs, ainsi que les intervalles de confiance à 95%, sont présentées figure 8.11 en fonction du débit de codage utilisé, pour chaque contenu et sur chacun des deux écrans. Les notes de qualité obtenues sur CRT sont significativement supérieures pour les débits les plus faibles, en revanche la différence entre les deux écrans n'est pas significative pour les séquences de référence non dégradées.

Le tableau 8.6 donne les valeurs de ΔQ_{MOY} et ΔQ_{REF} pour les quatre contenus de la base SVT. Les résultats sont donnés pour le format QHD ainsi que pour le format HD. Pour chacun des quatre contenus, la valeur de ΔQ_{MOY} diminue fortement à basse résolution. En moyenne, ΔQ_{MOY} est égal à 3.4 au format QHD contre 8.9 en HD. De même, ΔQ_{REF} est égale à 4.6 en moyenne au format HD tandis qu'il est négatif avec une moyenne de -1.5 au format QHD. Les versions de référence des séquences *Knightshields* et *Stockholm* ont été préférées sur LCD à basse résolution tandis qu'elles avaient largement été préférées sur CRT à haute résolution.

Ce résultat semble confirmer l'hypothèse selon laquelle le mouvement est la cause principale de la différence de qualité sur LCD. En effet, la vitesse de déplacement des objets présents sur l'écran est deux fois supérieure dans des conditions HD que dans des conditions QHD. Dans les deux cas,

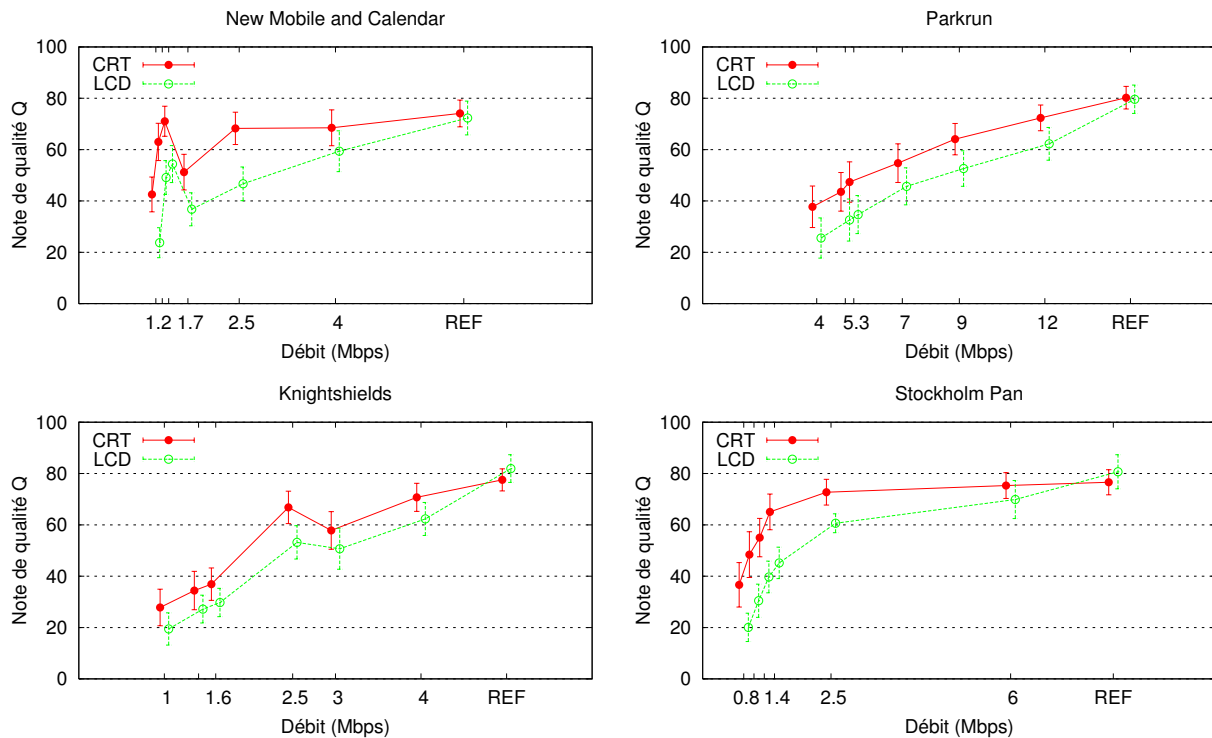


FIGURE 8.11 – Résultats de l'évaluation subjective des séquences SVT au format QHD, pour chaque débit et sur chaque type d'écran.

un degré visuel correspond à environ 60 pixels de l'écran, mais un objet se déplaçant à une vitesse de 10 pixels par image dans une séquence QHD (environ 8.3 deg/s), se déplacera à 20 pixels par image dans la même séquence au format HD (environ 16.7 deg/s). Cette augmentation de la vitesse de déplacement au niveau de l'œil implique un flou de mouvement plus important sur les écrans à cristaux liquides dans un contexte haute-définition. Cela se traduit donc par une perte de qualité plus importante par rapport à la qualité perçue sur un écran CRT.

En conclusion, nous avons montré avec cette série de tests supplémentaire que la diminution de la résolution des contenus implique une réduction la perte de qualité sur LCD. Les distorsions visuelles causées par le type d'affichage sont donc d'autant plus gênantes que la résolution de l'écran (et des contenus affichés) est importante. En particulier, le flou de mouvement est directement lié à l'amplitude des mouvements présents dans les séquences. L'augmentation de la résolution implique une augmentation de l'amplitude des mouvements et donc de la quantité de flou de mouvement perçue.

Séquence	ΔQ_{MOY}	ΔQ_{REF}	Séquence	ΔQ_{MOY}	ΔQ_{REF}
<i>New Mobcal</i> QHD	1.8	1.8	<i>New Mobcal</i> HD	10.0	4.0
<i>Parkrun</i> QHD	5.4	0.6	<i>Parkrun</i> HD	7.6	2.9
<i>Knightshields</i> QHD	2.0	-4.4	<i>Knightshields</i> HD	6.5	6.9
<i>Stockholm</i> QHD	4.5	-4.1	<i>Stockholm</i> HD	8.5	4.4
Moyenne QHD	3.4	-1.5	Moyenne HD	8.9	4.6

TABLE 8.6 – Différence de qualité entre CRT et LCD pour les quatre contenus de la base SVT. Comparaison entre les résultats obtenus au format QHD et ceux obtenus au format HD.

8.4 Conclusion

Dans ce chapitre, nous avons présenté les tests d'évaluation subjective de la qualité visuelle réalisés avec des séquences vidéos et des images fixes au format haute-définition. Ces tests se sont déroulés à la fois sur écran de type CRT et sur écran de type LCD. Ils ont permis de mettre en évidence une différence significative de qualité visuelle entre l'affichage sur écran CRT et l'affichage sur écran LCD. Une différence moyenne de 10.2 (sur une échelle de 1 à 100) en faveur du CRT a été mesurée sur l'ensemble des contenus vidéos, 6.1 en tenant compte seulement des séquences de référence non compressées. Pour la majorité des séquences, cette perte de qualité sur écran LCD semble être corrélée avec la quantité de mouvement. A l'inverse, les images fixes présentées aux observateurs ont obtenu des notes de qualité supérieures sur l'écran LCD.

Ces deux résultats confirment que la technologie d'affichage à cristaux liquides, adaptée à la présentation de contenus fixes, ne permet pas une reproduction de mouvement optimale. Plus précisément, les distorsions visuelles liées au mouvement, au premier rang desquelles le flou de mouvement, sont responsables d'une perte de qualité sur LCD. En outre, quatre séquences vidéos ont été testées à nouveau après avoir été redimensionnées à une résolution « quasi-standard » (960×540 pixels). La différence de qualité visuelle mesurée entre CRT et LCD est plus faible qu'au format HDTV (3.4 en moyenne au format standard contre 8.9 au format HDTV sur les quatre séquences considérées). Ce résultat confirme l'hypothèse selon laquelle cette différence est liée au mouvement. En effet, l'amplitude des mouvements est divisée par deux entre le format standard et le format haute-définition.

Ces résultats permettent donc de mettre en évidence trois points importants concernant la qualité visuelle sur les écrans à cristaux liquides :

- La qualité visuelle des séquences affichées sur un écran à cristaux liquides est réduite par les distorsions liées au mouvement.
- Plus la résolution d'affichage est élevée, plus ce phénomène est important. Négligeable pour des formats standard (TVSD), cette perte de qualité devient problématique pour des formats haute-définition (TVHD). Le flou de mouvement par exemple est deux fois plus important en haute-définition. Or, c'est le contexte HD qui donne tout leur intérêt aux écrans LCD (grands écrans pour profiter de l'augmentation des formats).
- D'un point de vue plus général, il convient donc de considérer l'influence de l'écran sur la qualité visuelle et plus seulement les distorsions classiques liées au codage, à la transmission, etc.

Dans le chapitre suivant, nous allons chercher à prédire cette différence de qualité visuelle entre

CRT et LCD. Cette prédiction est basée sur une modélisation du flou perçu global à partir des mouvements présents dans une séquence vidéo.

Chapitre 9

Vers une prédiction de la différence de qualité visuelle entre CRT et LCD basée sur le flou de mouvement

Sommaire

9.1	Introduction	161
9.2	Estimation locale du flou de mouvement perçu	162
9.2.1	Estimation de mouvement	163
9.2.2	Classification spatiale	164
9.2.3	Calcul du flou de mouvement	166
9.3	Prédiction de la différence de qualité visuelle entre CRT et LCD	167
9.3.1	Cumul des valeurs locales de flou de mouvement perçu	167
9.3.2	Prédiction de la différence de qualité ΔQ	170
9.3.3	Résultats	171
9.4	Observations et discussion	172
9.5	Conclusion	174

9.1 Introduction

Les tests d'évaluation subjective de la qualité visuelle présentés dans le chapitre précédent ont montré qu'une séquence vidéo obtient une note de qualité visuelle supérieure lorsqu'elle est affichée sur un écran CRT que lorsqu'elle est affichée sur un écran à cristaux liquides. Ce résultat est vrai pour des séquences originales, non compressées, qui ne contiennent intrinsèquement aucune dégradation. La différence de qualité est donc liée au dispositif d'affichage. De plus, ce phénomène est d'autant plus important en présence de forts mouvements, et lorsque la résolution visuelle des contenus affichés augmente. Ces résultats conduisent à l'hypothèse suivante : les distorsions d'écran liées au mouvement (flou de mouvement et désentrelacement principalement) sont les principaux responsables de la perte de qualité sur LCD.

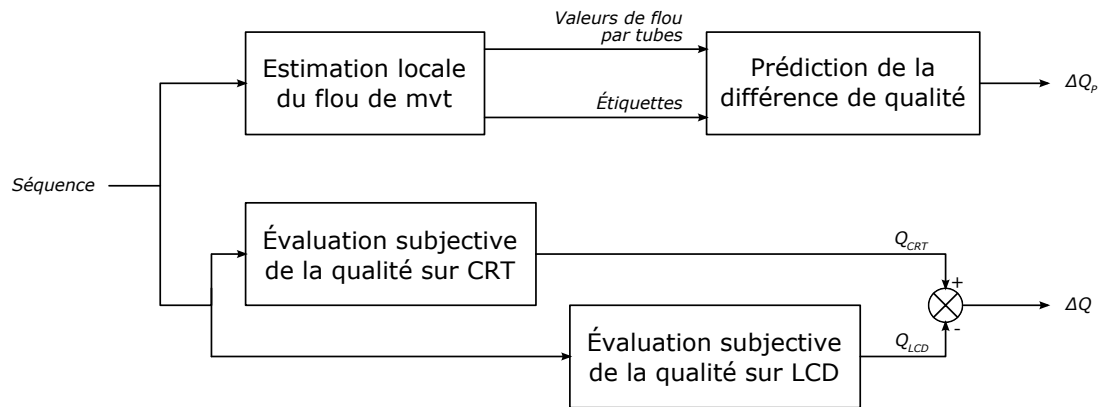


FIGURE 9.1 – Prédiction de la différence de qualité perçue entre écran CRT et écran LCD ΔQ_p .

L'objectif de ce chapitre est de combler le gap entre la caractérisation objective du flou de mouvement réalisée dans la partie 2 de ce mémoire, et l'évaluation de l'impact de cet artefact sur la qualité visuelle finale. Il est réalisé en deux parties (cf. figure 9.1). Tout d'abord, une estimation locale du flou de mouvement perçu (sur la séquence considérée) est mise en œuvre, à partir des caractéristiques mesurées de l'écran LCD et en se basant sur nos modèles psychovisuels. Puis la perte de qualité finale est ensuite prédite à partir des résultats de cette estimation. Cette prédiction automatique de la perte de qualité sur LCD nous permettra en effet de maîtriser l'intérêt d'une évolution technologique et d'évaluer son impact en terme de bénéfice sur la qualité perçue.

9.2 Estimation locale du flou de mouvement perçu

On l'a vu dans les chapitres précédents, le flou de mouvement apparaît sur les structures spatiales en déplacement (textures, contours) lorsque ces structures sont suivies des yeux par les observateurs. Le flou de mouvement perçu est alors proportionnel à la vitesse de déplacement des structures en mouvement. L'estimation du flou de mouvement local présent dans une séquence passe donc par deux étapes : identifier ces structures et connaître leur vitesse de déplacement. Une segmentation de la séquence vidéo en volumes spatio-temporels a donc été mise en œuvre. Elle s'appuie sur une estimation de mouvement basée bloc réalisée sur des tronçons de cinq images successives. Parallèlement, chaque volume est étiqueté en fonction de son contenu spatial (zone uniforme, textures faibles, textures fortes ou contours). Dans un second temps, le flou de mouvement perçu est estimé localement sur chaque volume à partir du vecteur de mouvement, et en tenant compte de la classification spatiale. La figure 9.2 illustre cette estimation.

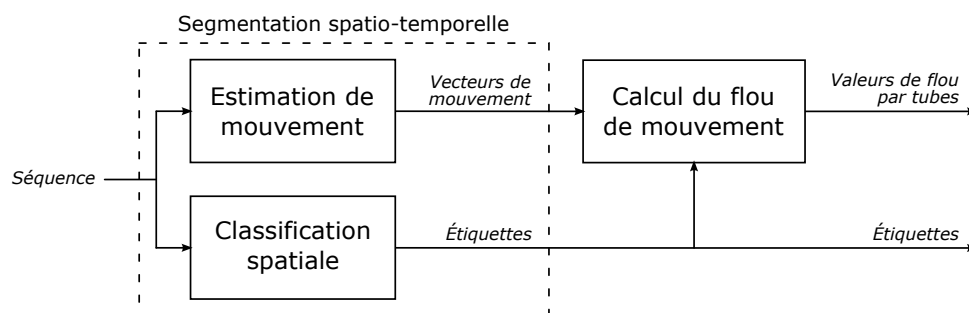


FIGURE 9.2 – Schéma-bloc de l'estimation locale du flou de mouvement perçu sur une séquence.

9.2.1 Estimation de mouvement

Les volumes spatio-temporels élémentaires, ou « tubes spatio-temporels », sont construits le long du mouvement local, sur des tronçons de cinq images successives (ou 10 trames dans le cas de séquences vidéo au format entrelacé). Cette durée élémentaire de 200 millisecondes a été choisie car c'est la durée moyenne d'une fixation entre deux saccades (cf. section A.3.3). La construction des tubes est réalisée via une estimation de mouvement basée bloc, sur chacun des tronçons de cinq images constituant la séquence. Ce concept de tubes tridimensionnels a été introduit par Wolf et Pinson [Wolf 99]. La nuance ici vient du fait que les tubes sont orientés temporellement selon la direction du mouvement et non pas dans une direction temporelle fixe.

La figure 9.3 illustre la construction des tubes spatio-temporels. L'image centrale du tronçon de cinq images, ou image « pivot », est découpée en blocs de taille 16×16 . Chaque bloc de l'image pivot est recherché simultanément dans les deux images précédentes et dans les deux images suivantes. Tous les vecteurs de mouvement correspondant à la zone de recherche sont évalués. Le vecteur de mouvement retenu est celui qui minimise l'erreur quadratique moyenne, calculée sur les trois composantes YCbCr, entre le bloc de l'image pivot et les quatre autres blocs. Cet ensemble de cinq blocs constitue un tube spatio-temporel.

Dans le cas de séquence vidéo entrelacé, cette construction de tubes spatio-temporels est effectuée indépendamment sur les cinq trames paires et sur les cinq trames impaires du tronçon, avec des blocs de taille 16×8 . Le vecteur de mouvement de chaque tube est alors calculé en moyennant les vecteurs mesurés sur les tubes de chaque parité. La complexité et le temps de calcul d'un tel traitement sur des séquences vidéo au format 1080i nous ont conduit à effectuer cette estimation de mouvement sur une représentation multi-résolution de chaque trame. Une première série de vecteurs est mesurée sur une représentation basse résolution du tronçon. Puis, les vecteurs de mouvement obtenus sont affinés à un niveau de résolution supérieur, et ainsi de suite. Trois niveaux de résolution ont été utilisés, permettant de réduire significativement le temps de calcul.

A l'issue de cette première étape, chaque séquence est découpée en tronçons en tubes spatio-temporels de taille $16 \times 16 \times 5$, auxquels est associé un vecteur de mouvement. Dans la suite, le tronçon courant est identifié par la variable k , et les coordonnées du tube dans l'image pivot par

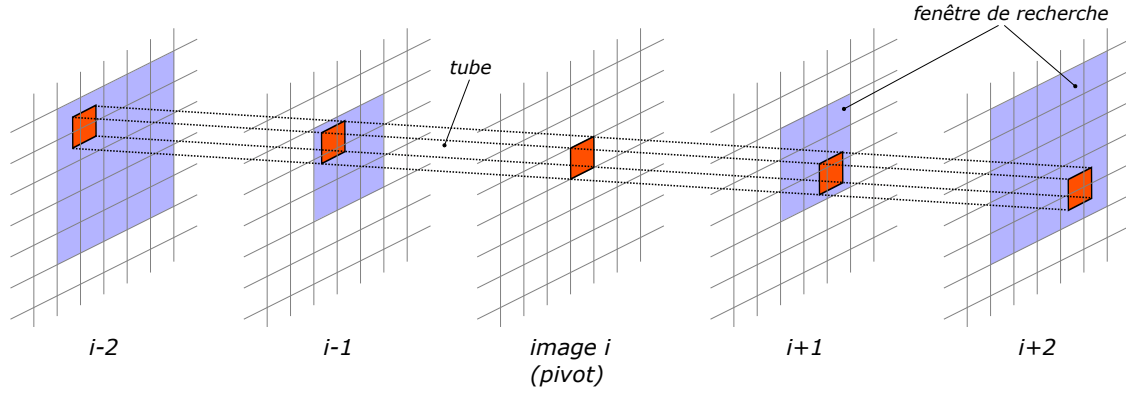


FIGURE 9.3 – Construction des tubes spatio-temporels par une estimation de mouvement par blocs de taille $m \times n$, sur un groupe de 5 images consécutives. L'estimation est en fait réalisée sur les trames paires et sur les trames impaires indépendamment (avec des blocs de taille $m \times \frac{n}{2}$). Les vecteurs de mouvement obtenus sur les tubes de chaque parité sont ensuite moyennés.

i et j . On identifiera donc un tube par la variable $T(k, i, j)$. Le vecteur de mouvement associé au tube $T(k, i, j)$ est noté $\vec{V}_{k,i,j}$, sa norme est notée $V_{k,i,j}$.

9.2.2 Classification spatiale

La seconde étape de cette segmentation spatio-temporelle consiste à étiqueter chaque tube selon son contenu spatial. Quatre classes, notées C_1 à C_4 , ont été définies selon l'activité spatiale de la zone considérée :

- Zones uniformes (C_1) : ce sont les zones pour lesquelles l'activité spatiale est très faible
- Textures faibles (C_2) : elles correspondent aux zones avec une activité spatiale réduite mais visible, par exemple des cheveux, une pelouse, etc.
- Textures fortes (C_3) : elles correspondent aux zones avec une activité spatiale forte sans orientation particulière, ou avec des orientations multiples, par exemple le feuillage des arbres.
- Contours (C_4) : les contours sont les zones d'activité spatiale forte dans lesquelles une orientation privilégiée est détectée.

Les caractéristiques d'activité spatiale de chaque tube sont mesurées par l'intermédiaire des gradient horizontal et vertical ΔH et ΔV de l'image pivot du tube spatio-temporel considéré. Dans un premier temps, la valeur $I_P(x, y)$ de chaque pixel de l'image pivot I_P est convertie en une valeur de luminance perceptuelle $I_P^*(x, y)$ selon la relation A.2. Les gradients sont ensuite calculés pour chaque pixel par les opérations de convolution bidimensionnelles suivantes :

$$\Delta H(x, y) = I_P^*(x, y) * S_H \quad (9.1)$$

$$\Delta V(x, y) = I_P^*(x, y) * S_V \quad (9.2)$$

avec S_H et S_V les filtres de Sobel horizontaux et verticaux de dimension 3×3 :

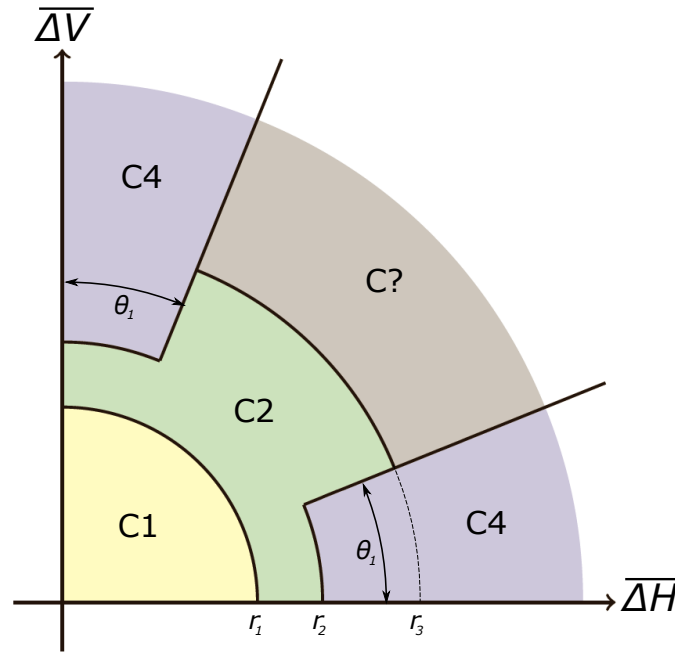


FIGURE 9.4 – Définition des classes par une partition de l'espace $(\Delta H, \Delta V)$. Dans la zone étiquetée C?, il y a ambiguïté entre les contours obliques (classe C4) et les textures fortes (classe C3).

$$S_H = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -2 \end{bmatrix} \quad S_V = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad (9.3)$$

Enfin, la moyenne des gradient est calculée sur chaque tube de coordonnées (i, j) :

$$\overline{\Delta H}(i, j) = \sum_{(x,y) \in \text{tube}(i,j)} \Delta H(x, y) \quad (9.4)$$

$$\overline{\Delta V}(i, j) = \sum_{(x,y) \in \text{tube}(i,j)} \Delta V(x, y) \quad (9.5)$$

Chaque tube peut à présent être étiqueté en fonction de la valeurs des gradients moyens $\overline{\Delta H}$ et $\overline{\Delta V}$. On utilise pour cela une partition de l'espace $(\overline{\Delta H}, \overline{\Delta V})$ telle que présentée figure 9.4. Cette partition est effectuée selon la norme r et l'orientation θ des gradients moyens :

$$r = \sqrt{\overline{\Delta H}^2 + \overline{\Delta V}^2} \quad (9.6)$$

$$\theta = \arctan\left(\frac{\overline{\Delta V}}{\overline{\Delta H}}\right) \quad (9.7)$$

Les valeurs r_1 , r_2 , r_3 et θ_1 ont été déterminées empiriquement sur l'ensemble des séquences de manière à obtenir la meilleure classification possible.

La zone étiquetée $C?$ correspond aux cas où la norme des gradients est élevée ($r > r_3$) et où les orientations verticales horizontales ne sont pas privilégiées ($\theta_1 < \theta < 90 - \theta_1$). Dans ce cas, il est impossible de décider si le tube contient des textures fortes (activité spatiale élevée sans orientation particulière ou avec orientations multiples) ou bien s'il contient des contours obliques. Pour lever cette ambiguïté, on calcule les gradients diagonaux de l'image $\Delta D45$ et $\Delta D135$ à l'aide des filtres de Sobel suivants :

$$S_{D45} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix} \quad S_{D135} = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} \quad (9.8)$$

Les gradients moyens $\overline{\Delta D45}$ et $\overline{\Delta D135}$ sont calculés comme précédemment et l'espace $(\overline{\Delta D45}, \overline{\Delta D135})$ est partitionné de la même manière que l'espace $(\overline{\Delta H}, \overline{\Delta V})$ avec les mêmes valeurs de paramètres (figure 9.4). Si la classification est toujours indéterminée dans ce nouvelle espace, on attribue alors la classe C3 (textures fortes) au tube considéré.

Cette deuxième étape achève la segmentation spatio-temporelle de la séquence. Chaque tube $T(k, i, j)$ appartient à une classe C_l ($l \in \{1; 2; 3; 4\}$).

9.2.3 Calcul du flou de mouvement

Le flou de mouvement perçu localement sur chaque tube spatio-temporel de la séquence est proportionnel à la vitesse de déplacement de la structure. Chaque tube est composé de $16 \times 16 \times 5$ pixels, chacun soumis à des niveaux de gris différents, et donc à différentes transitions temporelles. Calculer la largeur de flou de mouvement sur chaque pixel du tube n'est pas envisageable car il faudrait pour cela connaître la réponse indicielle des pixels pour chacune des 65280 transitions existantes. La valeur du flou de mouvement perçu sur l'ensemble du tube est donc estimée en utilisant le temps de flou global perceptif M_{PERC} défini dans le chapitre 6. Cette valeur est issue des mesures et des tests psychophysiques effectuées sur un ensemble de 20 transitions réparties sur l'ensemble de la dynamique lumineuse.

La largeur de flou de mouvement perçu BEW est donc calculé de la manière suivante pour chaque tube :

$$BEW = M_{PERC} \cdot V \quad (9.9)$$

où V est la norme du vecteur de mouvement du tube considéré.

Cependant, les tubes étiquetés comme zones uniformes (classe C_1) ne contiennent aucune structure spatiale (norme des gradients moyens inférieure à une valeur seuil r_1). Pour cette raison, le flou de mouvement perçu sur ces tubes est considéré comme nul. Finalement, le flou de mouvement perçu sur un tube de coordonnées (i, j) dans le tronçon k s'écrit :

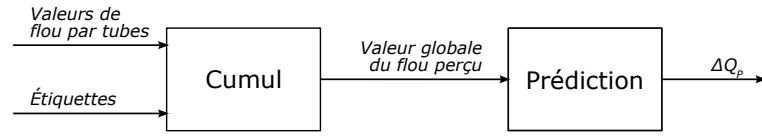


FIGURE 9.5 – Schéma-bloc de la prédiction de la différence de qualité entre CRT et LCD ΔQ_P à partir des valeurs locales de flou de mouvement perçu.

$$\begin{cases} BEW_{k,i,j} = 0 & \text{si } T(k,i,j) \in C_1 \\ BEW_{k,i,j} = M_{PERC} \cdot V_{k,i,j} & \text{sinon} \end{cases} \quad (9.10)$$

Cette première section a décrit l'estimation locale du flou de mouvement perçu dans une séquence. La séquence vidéo considérée est découpée en tronçons de cinq images, eux-mêmes découpés en tubes spatio-temporels de taille $16 \times 16 \times 5$. La valeur du flou de mouvement perçu est estimé sur chaque tube. Dans la section suivante, le cumul des valeurs locales du flou de mouvement perçu est mis en œuvre, avec pour objectif d'obtenir un indicateur global du flou de mouvement perçu qui permette de prédire la perte de qualité visuelle sur LCD.

9.3 Prédiction de la différence de qualité visuelle entre CRT et LCD

Nous cherchons à présent à prédire la différence de qualité visuelle entre CRT et LCD pour une séquence vidéo donnée. Cette prédiction est réalisée à partir des résultats de l'estimation locale du flou de mouvement perçu décrite dans la section précédente. La figure 9.5 présente les différentes étapes de l'opération. Dans un premier temps, les valeurs locales de flou de mouvement perçu sont cumulées pour obtenir une valeur globale du flou de mouvement perçu sur l'ensemble de la séquence. La seconde étape correspond à la prédiction proprement dite de la différence de qualité entre les deux types d'écran, à partir de cette valeur de flou globale.

9.3.1 Cumul des valeurs locales de flou de mouvement perçu

L'estimation locale du flou de mouvement présentée dans la section précédente a permis le découpage de la séquence en tubes spatio-temporels de taille $16 \times 16 \times 5$. Pour chaque tube, la valeur locale du flou de mouvement perçu a été calculée et une classification a été réalisée en fonction de l'activité spatiale. À partir de ces valeurs locales du flou de mouvement, une valeur globale de flou de mouvement perçu est calculée sur l'ensemble de la séquence. Quatre méthodes de cumul ont été mise en œuvre :

- en tenant compte de la classification spatiale, avec un cumul spatio-temporel sur l'ensemble de la séquence ;
- en tenant compte de la classification spatiale, avec un cumul spatio-temporel par tronçon, puis une moyenne sur l'ensemble des tronçons ;

- sans tenir compte de la classification spatiale, avec un cumul spatio-temporel sur l'ensemble de la séquence ;
- sans tenir compte de la classification spatiale, avec un cumul spatio-temporel par tronçon, puis une moyenne sur l'ensemble des tronçons.

Dans le premier cas les valeurs de flou calculées sur chaque classe possède le même poids, la valeur globale est ensuite calculée en pondérant ces trois valeurs. Dans le second cas, chaque tronçon possède la même pondération dans la valeur de flou associée à chaque classe. Ici encore, la valeur globale est calculée en pondérant les valeurs de flou calculées sur chacune des trois classes.

Dans les deux derniers cas, la classification spatiale n'est pas prise en compte, le flou perçu global est donc une combinaison des valeurs de flou perçu sur chaque classe, pondérées par la proportion de celles-ci sur l'ensemble de la séquence dans le premier cas, sur chaque tronçon dans le second cas.

9.3.1.1 Méthodes avec prise en compte de la classification spatiale

Dans ces deux premières méthodes, la classification spatiale de la séquence décrite dans la section précédente est prise en compte. Deux stratégies de cumul ont été utilisées. La première stratégie consiste à cumuler les valeurs locales de flou de mouvement perçu sur l'ensemble de la séquence, pour chaque classe. Avec la seconde stratégie, le cumul est d'abord effectué sur chaque tronçon, puis les valeurs obtenues par tronçon sont moyennées sur la séquence entière.

Cumul spatio-temporel sur l'ensemble de la séquence

La valeur globale du flou de mouvement perçu sur une classe donnée C_l ($l \in \{2; 3; 4\}$) est notée $BEW(C_l)$. Elle est calculée par la moyenne spatio-temporelle des valeurs locales de flou de mouvement perçu calculées sur l'ensemble des tubes étiquetés C_l dans la séquence. C'est à dire :

$$BEW(C_l) = \frac{1}{N_l} \sum_{\{k,i,j\} | T(k,i,j) \in C_l} BEW_{k,i,j} \quad (9.11)$$

avec N_l le nombre de tubes étiquetés C_l dans la séquence.

Cumul spatio-temporel par tronçon

Avec cette seconde stratégie, une valeur intermédiaire de flou de mouvement perçu est tout d'abord calculée sur chaque tronçon k , pour chaque classe C_l . Cette valeur intermédiaire est notée $BEW_k(C_l)$. Elle est calculée comme la moyenne spatiale des valeurs locales de flou de mouvement perçu calculées sur l'ensemble des tubes étiquetés C_l sur le tronçon k :

$$BEW_k(C_l) = \frac{1}{N_{l,k}} \sum_{\{i,j\} | T(k,i,j) \in C_l} BEW_{k,i,j} \quad (9.12)$$

avec $N_{l,k}$ le nombre de tubes étiquetés C_l dans le tronçon k .

La valeur globale du flou de mouvement perçu sur une classe donnée est ensuite calculée par la moyenne temporelle de l'ensemble de ces valeurs intermédiaires. Avec cette stratégie de cumul, chaque tronçon possède le même poids dans la valeur finale du flou de mouvement perçu associée à chaque classe. On a donc :

$$BEW(C_l) = \frac{1}{N_t} \sum_{k=1}^{N_t} BEW_k(C_l) \quad (9.13)$$

avec N_t le nombre de tronçons dans la séquence considérée.

Cumul des classes

Les deux stratégies de cumul considérées permettent de calculer une valeur globale de flou de mouvement perçu sur l'ensemble de la séquence, pour une classe donnée. Nous obtenons donc trois valeurs globales de flou de mouvement perçu $BEW(C_2)$, $BEW(C_3)$ et $BEW(C_4)$, correspondant respectivement aux classes C_2 (textures faibles), C_3 (textures fortes) et C_4 (contours).

Ces trois valeurs sont ensuite cumulées de manière à obtenir une valeur globale de flou perçu sur l'ensemble de la séquence BEW_G . Ce cumul des classes est réalisé par une combinaison linéaire des trois valeurs :

$$BEW_G = \frac{a \cdot BEW(C_2) + b \cdot BEW(C_3) + c \cdot BEW(C_4)}{a + b + c} \quad (9.14)$$

avec a , b et c trois nombres réels positifs inférieurs à 1.

Ces deux premières méthodes de cumul ont donc permis de calculer une valeur globale de flou de mouvement perçu sur l'ensemble de la séquence à partir des valeurs locales estimées sur chaque tube spatio-temporel. Ce cumul a été effectué en s'appuyant sur la classification spatiale des tubes, à travers deux stratégies de cumul spatio-temporel.

9.3.1.2 Méthodes sans prise en compte de la classification spatiale

Ces deux méthodes de cumul consistent à calculer la valeur globale de flou de mouvement perçu sur la séquence sans tenir compte de la classification spatiale. Cette valeur globale est donc le résultat d'un cumul sur l'ensemble de classes C_Σ qui correspond à l'union des classes C_2 , C_3 et C_4 :

$$C_\Sigma = C_2 \cup C_3 \cup C_4 \quad (9.15)$$

Deux stratégies de cumul ont été utilisées, les mêmes que précédemment : un cumul sur l'ensemble de la séquence et un cumul par tronçon. La pondération de chaque classe est alors liée à sa proportion dans la séquence ou dans le tronçon, selon la stratégie de cumul utilisée.

Cumul spatio-temporel sur l'ensemble de la séquence

La valeur globale de flou de mouvement perçu sur la séquence BEW_G est calculée comme la moyenne des valeurs locales de flou perçu estimées sur l'ensemble des tubes :

$$BEW_G = \frac{1}{N_\Sigma} \sum_{\{k,i,j\} | T(k,i,j) \in C_\Sigma} BEW_{k,i,j} \quad (9.16)$$

avec N_Σ le nombre de tubes étiquetés C_2 , C_3 ou C_4 dans la séquence.

Cumul spatio-temporel par tronçon

Dans ce cas là, une valeur intermédiaire de flou de mouvement perçu est calculée sur chaque tronçon k , pour l'ensemble C_Σ :

$$BEW_k = \frac{1}{N_{\Sigma,k}} \sum_{\{i,j\} | T(k,i,j) \in C_\Sigma} BEW_{k,i,j} \quad (9.17)$$

avec $N_{\Sigma,k}$ le nombre de tubes appartenant à l'ensemble C_Σ

La valeur globale de flou de mouvement perçu BEW_G est alors obtenue en faisant la moyenne sur l'ensemble de la séquence des valeurs obtenues pour chaque tronçon :

$$BEW_G = \frac{1}{N_t} \sum_{k=1}^{N_t} BEW_k \quad (9.18)$$

avec N_t le nombre de tronçons dans la séquence considérée.

Ces deux méthodes de cumul ont permis de calculer une valeur globale de flou de mouvement perçu sur l'ensemble de la séquence à partir des valeurs locales estimées sur chaque tube spatio-temporel. Ce cumul a été effectué sur l'ensemble des tubes sans tenir compte de la classification spatiale de ces tubes.

9.3.2 Prédiction de la différence de qualité ΔQ

La prédiction de la différence de qualité perçue entre CRT et LCD ΔQ est effectuée à partir du flou de mouvement perçu global BEW_G calculé précédemment. Une fonction affine est utilisée pour obtenir une prédiction ΔQ_P de cette perte de qualité sur LCD :

$$\Delta Q_P = d \cdot BEW_G + e \quad (9.19)$$

avec d et e deux nombres réels, la valeur BEW_G est exprimée en minute d'arc.

Quatre méthodes de cumul ont été définies précédemment, on obtient par conséquent quatre modèles de prédictions en associant à ces quatre méthodes de cumul l'équation précédente. Les deux premiers modèles tiennent compte de la classification spatiale des tubes spatio-temporels, et possède cinq paramètres a , b , c , d et e présents dans les équations 9.14 et 9.19. Les deux autres modèles ne tiennent pas compte de cette classification et possède seulement deux paramètres d et e (équation 9.19).

Pour les quatre modèles de prédiction, les paramètres ont été optimisés par la méthode des moindres carrés de manière à obtenir la meilleure prédiction ΔQ_P des données expérimentales ΔQ_{REF} obte-

nues dans le chapitre précédent. Cette prédiction de la différence de qualité entre CRT et LCD est basée sur la quantité de flou de mouvement perçu dans la séquence. Elle s'appuie sur l'hypothèse selon laquelle le flou de mouvement est responsable de la perte de qualité visuelle mesurée sur LCD. Pour cette raison, l'optimisation des paramètres a été effectuée sur les séquences pour lesquelles la valeur globale de flou de mouvement perçu est supérieure au seuil de perception BEW_p^{perc} mesuré dans le chapitre 7 ($BEW_p^{perc} = 4.0$ minutes d'arc). Dans le cas contraire, on peut supposer que le flou de mouvement n'a pas eu d'influence sur le jugement des observateurs.

Sur les séquences *Movie* et *Standing*, le flou de mouvement perçu mesuré est inférieur à BEW_p^{perc} . Ces deux séquences n'ont donc pas été prises en compte dans la construction du modèle de prédiction et dans l'optimisation des paramètres. De la même, la séquence *Golf* a obtenu une note de qualité subjective supérieure sur l'écran LCD que sur l'écran CRT. Le flou de mouvement mesuré sur cette séquence est assez faible également. Cette séquence n'a donc pas été prise en compte non plus, dans la mesure où il semblerait que d'autres facteurs ont joué un rôle plus déterminant que le flou de mouvement.

Au final, la construction des modèles de prédiction de la différence de qualité entre CRT et LCD, via l'optimisation des paramètres a , b , c , d et e , a été réalisée sur les données subjectives obtenues sur neuf séquences de référence. Deux modèles ont été mis en œuvre, le premier tient compte de la classification spatiale de la séquence, le second non. Pour chacun des deux modèles, deux stratégies de cumul ont été testées. L'optimisation des paramètres a été réalisée pour chacune de ces quatre conditions.

9.3.3 Résultats

Le tableau 9.1 présente les cinq paramètres obtenus après optimisation par la méthode des moindres carrés, dans le cas des deux modèles de prédiction avec prise en compte de la classification. Le coefficient de corrélation linéaire (CCL) et la racine carrée de l'erreur quadratique moyenne (RMSE) entre la différence de qualité prédite ΔQ_P et la différence de qualité mesurée subjectivement ΔQ_{REF} sont également indiqués. Une optimisation sur les deux paramètres d et e a également été effectuée dans le cas où les paramètres a , b et c sont fixés à 1. Cette condition correspond à une pondération identique sur les trois classes spatiales.

Dans le cas du cumul spatio-temporel sur la séquence entière, on obtient une pondération de 1 sur les classes C_3 (textures fortes) et C_4 (contours) et une pondération de 0.75 sur la classe C_2 (textures faibles). Cependant, le résultat est sensiblement identique sans pondération en regard des indicateurs que sont le CCL et la RMSE. Ce résultat est cependant assez conforme à ce que l'on pourrait attendre. Les contours et les textures fortes étant les zones sur lesquelles le flou de mouvement est le plus visible.

Dans le cas du cumul par tronçon, la pondération obtenue est de 1 sur les contours, de 0.97 sur les textures faibles et de 0.17 sur les textures fortes. Cette pondération est en contradiction avec la précédente. Néanmoins, la qualité de la prédiction est plus forte avec un CCL de 0.9662 et une RMSE de 1.04. Avec cette stratégie de cumul, le résultat sans pondération est très différent, de moindre qualité.

Cumul spatio-temporel	a	b	c	d	e	CCL	RMSE
sur la séquence entière	0.75	1.00	1.00	1.66	-11.0	0.9384	1.39
”	<i>1</i>	<i>1</i>	<i>1</i>	<i>1.59</i>	<i>-9.86</i>	<i>0.9364</i>	<i>1.41</i>
sur chaque tronçon	0.97	0.17	1.00	1.74	-12.5	0.9662	1.04
”	<i>1</i>	<i>1</i>	<i>1</i>	<i>1.49</i>	<i>-8.89</i>	<i>0.8900</i>	<i>1.83</i>

TABLE 9.1 – Valeurs optimales des cinq paramètres du modèle de prédiction de la différence de qualité entre CRT et LCD avec prise en compte de la classification spatiale. Le coefficient de corrélation linéaire (CCL) et la racine carrée de l’erreur quadratique moyenne (RMSE) sont également précisés. Les résultats obtenus dans le cas d’une pondération uniforme des trois classes sont présentés en italique.

Cumul spatio-temporel	d	e	CCL	RMSE
sur la séquence entière	1.65	-10.6	0.9471	1.29
sur chaque tronçon	1.71	-11.6	0.9697	0.98

TABLE 9.2 – Valeurs optimales des deux paramètres du modèle de prédiction de la différence de qualité entre CRT et LCD sans prise en compte de la classification spatiale. Le coefficient de corrélation linéaire (CCL) et la racine carrée de l’erreur quadratique moyenne (RMSE) sont également précisés.

Le tableau 9.2 présente les deux paramètres obtenus après optimisation par la méthode des moindres carrés, dans le cas des deux modèles de prédiction sans prise en compte de la classification spatiale. Ici encore, la meilleur prédiction est obtenue avec le cumul par tronçon. Les résultats conduisent à un CCL de 0.9697 et à une RMSE de 0.98 entre les valeurs prédites ΔQ_P et les valeurs réelles ΔQ_{REF} .

9.4 Observations et discussion

La prédiction de la différence de qualité visuelle entre CRT et LCD à partir des valeurs de flou de mouvement perçu estimées localement donne des résultats satisfaisants avec les quatre méthodes de cumul. Néanmoins, c’est la méthode qui ne tient pas compte de la classification spatiale et qui utilise un cumul spatio-temporel par tronçons qui obtient les meilleures performances.

Avec cette méthode de cumul, les valeurs de flou perçu sur chaque classe sont pondérées selon la proportion de la classe dans le tronçon courant. La classification entre cependant en jeu dans le fait que les tubes étiquetés C_1 (zones uniformes) ne sont pas pris en compte. Chaque tronçon est ensuite affecté du même poids temporel dans le calcul du flou de mouvement perçu global sur l’ensemble de la séquence. La figure 9.6 présente la différence de qualité entre CRT et LCD mesurée subjectivement ΔQ_{REF} en fonction du flou de mouvement perçu global calculé avec cette méthode de cumul $BEW_p^{gène}$, pour les douze séquences vidéo. La ligne pleine représente la prédiction de cette différence ΔQ_P . Le seuil de perception BEW_p^{perc} , le seuil de gêne $BEW_p^{gène}$ et le seuil d’acceptabilité

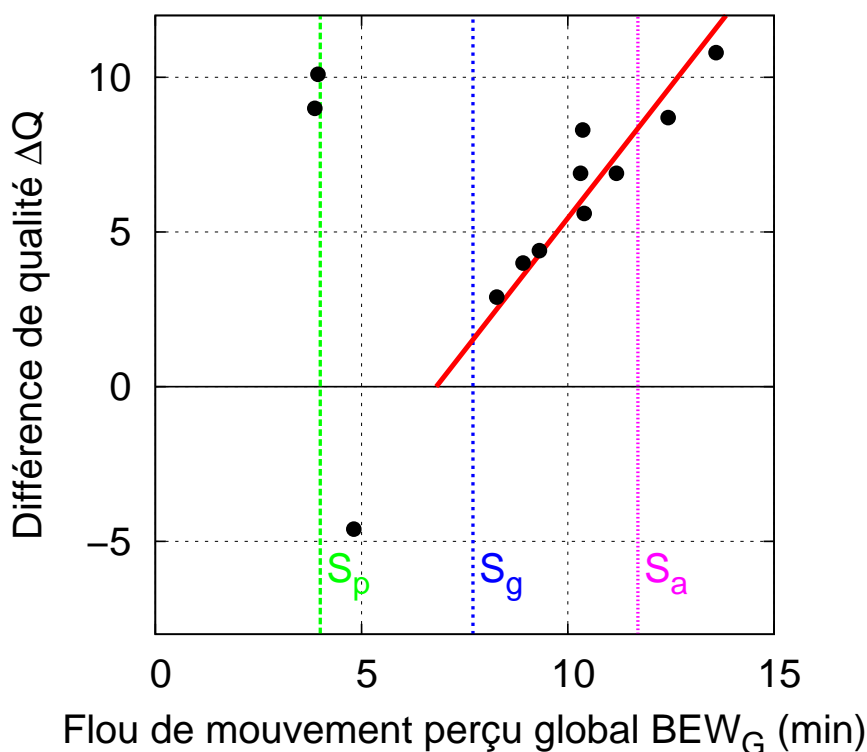


FIGURE 9.6 – Les points représentent la différence de qualité visuelle mesurée subjectivement entre CRT et LCD (ΔQ_{REF}) en fonction du flou de mouvement perçu global (BEW_G) exprimé en minutes d’arc. Les trois seuils de flou S_p , S_g et S_a (respectivement seuil de perception, seuil de gêne et seuil d’acceptabilité) mesurés dans le chapitre 7 sont indiqués sur la figure. La ligne pleine représente la prédiction ΔQ_P (équation 9.19).

BEW_p^{accept} , déterminés dans le chapitre 7, sont représentés par les lignes verticales pointillées sur la figure 9.6.

Pour les neuf séquences vidéos présentant une valeur globale de flou de mouvement perçu supérieure au seuil de gêne S_g , la différence de qualité visuelle entre CRT et LCD s’accroît au fur et à mesure que la quantité de flou augmente. On remarque qu’à partir de ce seuil de gêne (environ 7 minutes d’arc), le flou de mouvement semble être la contribution principale à la perte de qualité sur LCD. La différence de qualité augmente d’environ 1.7 points (sur une échelle de 100) pour chaque minute d’arc.

En deçà du seuil de gêne, le flou de mouvement perçu sur la séquence n’est plus l’élément principal sur lequel se base le jugement des observateurs. D’autres facteurs, toujours liés aux différences entre les deux technologies d’affichage, entrent alors en jeu. Trois séquences vidéo présentent une valeur globale de flou de mouvement perçu inférieure au seuil de gêne, proche du seuil de perception. Pour les séquences *Movie* et *Standing*, la différence de qualité est élevée malgré un flou de mouvement très faible. Ces deux séquences sont assez sombres et un fort bruit d’acquisition est présent. Ce

bruit est très visible sur LCD à cause de la dynamique de luminance plus élevée de cet écran, et à cause d'une reproduction des noirs non optimale sur les écrans à cristaux à liquides (luminance des noirs élevée par rapport aux écrans de type CRT). La séquence *Golf* est la seule séquence dont la visualisation a été préférée sur LCD par les observateurs ($\Delta Q_{REF} < 0$). Cette séquence est très lumineuse, avec des couleurs saturées et peu de flou de mouvement. Dans ces conditions, l'écran LCD, avec une gamme de luminance environ 6 fois plus élevée que l'écran CRT, affiche une image plus belle, tandis qu'elle paraît assez terne sur l'écran CRT.

En d'autres termes, les résultats de cette étude semblent montrer que le flou de mouvement est la principale contribution à la perte de qualité visuelle sur les écrans à cristaux liquides, lorsqu'il dépasse un seuil de gêne de l'ordre de 7 minutes d'arc. En deçà de ce seuil de gêne, les autres différences entre les deux types d'écrans contribuent au jugement des observateurs. Les séquences lumineuses avec beaucoup de couleurs sont préférées sur LCD, on retrouve alors les résultats des tests d'évaluation de la qualité effectués sur des images fixes (cf. section 8.3.2). Cependant les séquences plus sombres mettent en défaut la reproduction du noir sur les écrans à cristaux liquides et sont préférées sur CRT. Ce défaut est cependant de moins en moins important sur les dernières générations d'écrans LCD et notamment sur les écrans équipés d'un rétro-éclairage à DEL.

9.5 Conclusion

Dans ce chapitre, une estimation du flou de mouvement perçu par l'observateur a été réalisée localement, par l'intermédiaire d'une segmentation spatio-temporelle de la séquence. Cette segmentation basée sur une estimation de mouvement a permis le découpage de la séquence en tubes spatio-temporels auxquels sont associés des vecteurs de mouvement. Le flou de mouvement perçu est alors estimé sur chaque tube à l'aide du temps de flou perceptif mesuré sur l'écran dans le chapitre 6. Dans un second temps, une valeur globale du flou de mouvement perçu a été calculée sur l'ensemble de la séquence en cumulant les valeurs de flou estimée sur chaque tube. Ce cumul est réalisé sur l'ensemble des tubes contenant des textures et des contours. Les tubes ne contenant que des zones uniformes ne sont pas considérés.

Nous avons montré ensuite que la différence de qualité entre CRT et LCD, ou plus précisément la perte de qualité visuelle observée sur les écrans LCD dans le chapitre précédent pouvait être prédite à partir de la valeur globale de flou de mouvement perçu sur la séquence. En effet, au delà d'une valeur de flou perçu de l'ordre de 7 minutes d'arc (identifiée comme le seuil de gêne S_g déterminé dans le chapitre 7) la perte de qualité sur LCD s'accroît à mesure que le flou augmente. En deçà de cette valeur seuil, le flou de mouvement n'est plus gênant et d'autres facteurs entrent alors en jeu.

Le modèle de prédiction de la différence de qualité entre CRT et LCD présenté dans ce chapitre n'est pas complet en ce sens qu'il ne permet que la prédiction de la perte de qualité sur LCD au delà d'un certain seuil de flou de mouvement perçu. Néanmoins, cette étude a permis de montrer que le flou de mouvement, lorsqu'il dépasse le seuil de gêne, est le facteur prépondérant qui guide le jugement des observateurs. Elle a également permis de confirmer que la contribution du flou de mouvement à la perte de qualité est négligeable en deçà d'un certain seuil. D'après le modèle de prédiction, sur un écran de qualité excellente, avec un temps de flou global perceptif inférieur à 6.5

ms (cf. tableau 7.9), la contribution du flou de mouvement à la différence de qualité aurait bien été nulle.

Conclusion et perspectives

Les travaux présentés dans ce mémoire de thèse s’articulaient autour de trois objectifs : la caractérisation objective du flou de mouvement par un protocole de mesures photométriques, la modélisation de la perception du flou de mouvement à travers un ensemble d’expérimentations psychovisuelle, et enfin l’étude de l’impact du flou de mouvement sur la qualité perçue. Dans cette conclusion générale, nous présentons une synthèse de ces travaux en insistant sur les contributions originales dont ils sont responsables. Ce bilan est suivi par un ensemble de perspectives concernant chacun des trois axes de l’étude.

Synthèse des travaux réalisés

Caractérisation objective du flou de mouvement

Ces travaux ont été réalisés au sein du laboratoire Acreo AB, dans le cadre d’un projet de recherche piloté par l’organisme de certification TCO Development ¹. Nos contributions à la caractérisation objective du flou de mouvement sont multiples.

Dans un premier temps, nous avons développé un protocole de mesure du flou de mouvement simple et rapide, basé sur la réponse temporelle indicielle des pixels de l’écran. Ce protocole de mesure original trouve son origine dans l’analyse de la formation du flou de mouvement sur la rétine. Cette analyse montre que l’on peut calculer le profil spatial flou d’un contour en déplacement par une convolution de la réponse temporelle indicielle des pixels de l’écran avec une fonction porte. Le principal avantage de cette méthode réside dans sa simplicité de mise en œuvre : il suffit d’un simple capteur photo-électrique (avec une fréquence d’échantillonnage suffisante) pour obtenir une mesure précise du flou de mouvement, pour n’importe quelle vitesse de déplacement. Cela permet donc de s’affranchir des problèmes de calibrage, de synchronisation, et plus généralement de reproductibilité, des méthodes actuelles basées sur l’utilisation d’un dispositif imageur.

Dans un second temps, nous avons évalué cette méthode analytique de mesure du flou de mouvement en regard des méthodes existantes. Une campagne de mesures d’un ensemble de 20 transitions gris-à-gris sur quatre écrans différents a été menée en utilisant à la fois notre protocole et un protocole « classique » (utilisation d’une caméra rapide). Les deux protocoles de mesure ont donné des résultats équivalents, sauf dans certains cas particuliers où les avantages et défauts de chacune des deux méthodes ont pu être analysés. Cette campagne de mesure a également permis de montrer que le

1. <http://www.tcodevelopment.com>

temps de réponse n'était pas une mesure fiable pour évaluer le flou de mouvement, particulièrement pour les écrans à cristaux liquides équipés de dispositifs de réduction du flou de mouvement.

Ce protocole de mesure du flou de mouvement basé sur la réponse indicielle temporelle des pixels de l'écran permet de mesurer simplement et rapidement le flou de mouvement d'un large ensemble de transitions à la fois. Il est particulièrement adapté à une utilisation par un ensemble de laboratoires car la mesure de la réponse indicielle par un capteur photo-électrique est beaucoup moins sujette à variations qu'une mesure utilisant une caméra. Pour ces raisons, notre méthode, devrait être utilisée dans le cadre de la nouvelle version du label TCO Media Displays, qui doit intégrer une exigence concernant le flou de mouvement sur les écrans à cristaux liquides.

L'ensemble de ces travaux de caractérisation objective du flou de mouvement sur les écrans à cristaux liquides a fait l'objet d'une présentation lors de la conférence internationale *SID'2008*, puis une version étendue a été publiée dans la revue spécialisée *Journal of the Society for Information Display*.

Modélisation psychovisuelle du flou de mouvement

Nos travaux ont également donné lieu à plusieurs contributions concernant la modélisation psychovisuelle du flou de mouvement sur les écrans à cristaux liquides.

L'étude de la perception du flou de mouvement, grâce à un protocole d'égalisation de flou entre le flou perçu sur un contour en mouvement et un flou simulé à partir des mesures objectives, a mis en évidence une différence entre le flou de mouvement perçu par les observateurs et le flou mesuré de manière photométrique. Si le flou perçu est sensiblement égal au flou mesuré sur les contours contrastés, il devient plus faible que le flou mesuré lorsqu'on considère des contours moins contrastés. La pondération, en fonction de la luminance perçue L^* du contour considéré, des valeurs de flou de mouvement mesurées sur un ensemble donné de transitions, permet d'obtenir une valeur globale de flou de mouvement perceptuellement uniforme. Comparé à la simple moyenne des mesures, cette moyenne pondérée par la visibilité du flou permet d'estimer plus fidèlement le flou de mouvement perçu par les observateurs sur un écran donné. Ces travaux ont été présentés lors de la conférence internationale *Human Vision and Electronic Imaging 2009*.

Une seconde série d'expérimentations psychophysiques a été menée pour étudier la gêne provoquée par le flou de mouvement. En utilisant plusieurs types de stimuli (séquences synthétiques, textes défilants et images naturelles) sur lesquels la quantité de flou de mouvement perçu était connue a priori, nous avons pu mettre en relation la perception du flou et la sensation de gêne associée. Les résultats de ces expériences ont permis de définir les seuils de perception (4 arc min), de gêne (7.7 arc min) et de non acceptabilité (11.7 arc min) du flou de mouvement. À partir de ces valeurs seuils, et en tenant compte du mouvement moyen présent dans les signaux multimédia affichés, nous avons émis des recommandations concernant la valeur globale perceptive du temps de flou. Avec une valeur de temps de flou inférieure à 20 ms, la qualité de l'écran en termes de reproduction de mouvement est considérée acceptable. Avec une valeur inférieure à 11 ms, la qualité est jugée bonne. Enfin, la qualité de reproduction du mouvement sera excellente pour une valeur de temps de flou inférieure à 6.5 ms (flou de mouvement non perceptible dans la majorité des cas). Ces résultats ont

fait l'objet d'une présentation lors de la conférence internationale *SID'2009*.

Enfin, ces expérimentations psychophysiques ont également été l'occasion d'étudier un aspect précis de la vision humaine : l'influence du mouvement sur la réponse supraliminaire du système visuel. À travers une expérience originale d'égalisation de contraste entre un stimulus fixe et un stimulus en mouvement, nous avons pu construire une fonction de sensibilité supraliminaire aux contrastes en mouvement. Nous avons montré à travers ces résultats que le phénomène de « stabilisation du contraste » (*contrast constancy*) qui existe à des niveaux supraliminaires pour des stimuli purement spatiaux et pour des stimuli clignotants, existe également pour des stimuli en mouvement. Bien que ces résultats n'aient pas été utilisés pour caractériser la perception du flou de mouvement, ils constituent une contribution importante dans le domaine de la vision appliquée. De plus, l'étude de cette fonction de sensibilité supraliminaire nous a permis de prédire qualitativement que le flou de mouvement perçu serait moins important que le flou de mouvement mesuré. Cette étude psychophysique a été présentée lors de la conférence internationale *Human Vision and Electronic Imaging 2007*.

Impact du flou de mouvement sur la qualité visuelle

La campagne d'évaluation subjective de la qualité visuelle menées durant nos travaux de thèse a permis de souligner plusieurs points importants concernant l'évaluation de la qualité dans son ensemble. Depuis quelques années, l'évaluation subjective et objective de la qualité est devenu un axe de recherche à part entière, comme en témoigne par exemple les travaux du *Video Quality Expert Group* (VQEG). Cependant, les études passées et en cours concernent principalement l'évaluation de la qualité dans des contextes de codage ou de transmission. Le dispositif d'affichage est généralement considéré comme « transparent », en ce sens qu'il n'apporte pas de dégradations significatives en regard des autres distorsions. Avec l'arrivée des écrans à cristaux liquides et l'arrêt progressif de la production des écrans CRT, la question de la qualité des écrans utilisés pour les tests subjectifs commence à être posée. Dans ce contexte, notre étude a permis de montrer que la qualité visuelle, des séquences vidéos comme des images fixes, variait fortement selon que le contenu était affiché sur un écran ou sur un autre. Plus précisément, les séquences vidéos évaluées par un ensemble d'observateurs ont obtenues des notes subjectives très supérieures sur CRT que sur LCD (avec une différence de l'ordre de 10 points sur une échelle de 100). Dans un contexte d'évaluation objective de la qualité, l'écran doit donc être considéré dans la chaîne des traitements subis par le signal, au même titre que les étapes de transmission et/ou de compression.

De manière à répondre à ce problème, il est donc nécessaire de pouvoir modéliser l'impact de l'écran sur la qualité finale. Finalement, nous avons montré qu'une estimation locale du flou de mouvement, basée sur la modélisation psychovisuelle réalisée précédemment, pouvait conduire à une estimation fiable de la perte de qualité observée sur les écrans LCD. En effet, nous avons montré qu'au delà du seuil de gêne (7.7 arc min) la perte de qualité observée sur LCD était fortement corrélée avec la quantité de flou de mouvement perçu dans la séquence considérée. Ce résultat confirme que le flou de mouvement est le défaut le plus gênant sur les écrans à cristaux liquides lorsque le contenu affiché présente de forts mouvements. De plus, à l'heure où différentes solutions techniques sont

mise en œuvre par les fabricants pour réduire cet artefact, ce modèle de l'écran permet d'évaluer l'intérêt d'une évolution technologique en terme de bénéfice sur la qualité perçue, en tenant compte de manière automatique du couple écran-observateur.

Ces travaux portant sur l'évaluation de la qualité perçue ont fait l'objet de plusieurs présentations dans des conférences internationales (notamment *ICIP'2007* et *EUSIPCO'2007*). Finalement, une synthèse de ces présentations en conférences a été publiée dans la revue *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*.

Perspectives

Plusieurs perspectives peuvent être envisagées dans le prolongement de ces travaux de thèse.

La première concerne la modélisation du profil spatial d'un contour en mouvement sur l'ensemble des transitions gris-à-gris d'un écran. La modélisation de la réponse indicielle sur l'ensemble des transitions à partir de quelques réponses mesurées a été développée dans les travaux de thèse de Pierre Adam [Adam 08]. Grâce à notre méthode de mesure analytique basée sur la réponse indicielle des pixels de l'écran, il pourrait être possible d'étendre cette modélisation à la mesure du flou de mouvement. Cependant, la modélisation d'Adam ne tient pas compte des artefacts temporels qui peuvent exister sur la réponse indicielle, notamment ceux introduits par les dispositifs de réduction du flou (*backlight flashing*). Une solution pourrait être d'étendre cette modélisation à tout type de réponse indicielle, en traitant par exemple séparément les composantes basses-fréquences et hautes-fréquences du signal. Une telle modélisation permettrait de considérer un large ensemble de transitions gris-à-gris, et d'obtenir au final une valeur globale de flou de mouvement perçu beaucoup plus précise. En s'inspirant des travaux d'Adam, une autre solution pourrait être d'estimer directement le profil spatial d'un contour flou pour une transition quelconque à partir des mesures de flou réalisées sur quelques transitions.

Il pourrait également être intéressant d'étendre nos modèles à la couleur. La prise en compte de la couleur ne change rien en terme de mesure : les sous-pixels associés à chaque filtre chromatique sont semblables et leurs réponses indicelles sont les mêmes. Cependant, le flou de mouvement perçu sur un contour entre deux zones colorées peut s'avérer beaucoup plus visible et gênant qu'un flou en niveau de gris. En effet, dans ce cas ce n'est pas une seule transition mais trois transitions temporelles qui sont mises en jeu. Chacune de ces trois réponses indicelles provoque la formation d'un flou de mouvement plus ou moins large sur les trois composante chromatique de l'image. Les différences de largeur de ces trois zones floues provoque un déséquilibre dans le mélange de couleur au niveau du contour, pouvant conduire à la présence des couleurs très éloignées des deux couleurs initiales, et accentuer la visibilité du flou de mouvement et la gêne ressentie par l'observateur.

Enfin, dans l'optique de développer un modèle complet de l'écran à cristaux liquides, il serait intéressant d'étudier les autres défauts de l'écran LCD. À cet égard, la méthodologie globale développée dans ces travaux de thèse pour l'étude du flou de mouvement pourrait être appliquée à d'autres distorsions telles que l'affichage des zones sombres ou la reproduction des couleurs. De telles études permettrait d'étendre notre modèle de prédiction de la perte de qualité sur LCD à l'ensemble du

système dégradant que constitue l'écran.

Annexes

Annexe A

Propriétés utiles du système visuel humain

Sommaire

A.1	Introduction	185
A.2	Physiologie du système visuel humain	186
A.2.1	L'œil	186
A.2.2	La rétine	187
A.3	Adaptation du système visuel humain à son environnement	188
A.3.1	Sensibilité et adaptation à la lumière	189
A.3.2	Perception des couleurs	191
A.3.3	Mouvements de l'œil	194
A.4	Limitations du système visuel humain	195
A.4.1	Résolution spatiale et acuité visuelle	195
A.4.2	Résolution temporelle	201
A.5	Sensibilité au contraste et aux textures	203
A.5.1	Loi Weber-Fechner	203
A.5.2	Fonctions de sensibilité aux contrastes (CSF)	204
A.6	Perception du mouvement	207
A.6.1	Généralités	207
A.6.2	Cinéma et télévision : l'illusion du mouvement	208
A.7	Conclusion	210

A.1 Introduction

Dans cette annexe, les propriétés importantes du système visuel humain (SVH) sont abordées. La physiologie de l'œil est tout d'abord esquissée de façon sommaire. L'œil est un organe complexe constitué de différentes parties. La partie optique permet de former l'image de notre environnement et de la projeter sur la rétine. Cette dernière est en quelque sorte le capteur qui enregistre l'image

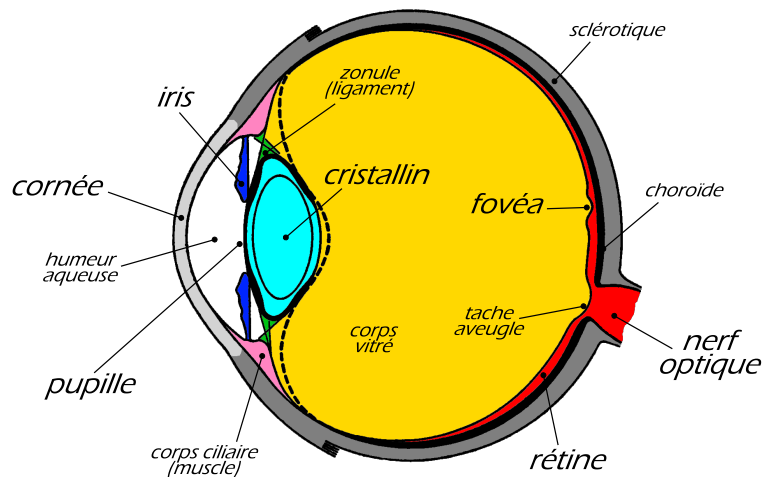


FIGURE A.1 – Anatomie de l'œil : vue en coupe de l'œil droit suivant une section horizontale.

et la transmet aux cellules neuronales. Enfin depuis les couches supérieures de la rétine et jusqu'au cortex visuel, l'information est traitée de manière à créer une représentation de notre environnement. L'œil a également su s'adapter à son environnement grâce à de nombreuses caractéristiques telles que les mouvement oculaires, l'adaptation au niveau de luminosité et la perception des couleurs. Bien entendu, le SVH est loin d'être parfait et possède des limitations intrinsèques qui réduisent la perception. Ces limitations interviennent principalement au niveau du système optique et de la rétine et sont décrites par la suite. La sensibilité au contraste est ensuite décrite plus précisément car c'est une propriété du SVH qui s'avère utile lorsqu'on étudie la perception d'une image affichée sur un écran. Enfin, nous décrirons succinctement la perception du mouvement et nous verrons comment les limitations du système visuel peuvent être utilisées pour abuser l'observateur, notamment au cinéma.

Les fonctionnement et les propriétés du système visuel humain ont fait l'objet de nombreux ouvrages et nous n'en décrivons ici que les concepts en relation avec l'affichage. Pour plus d'informations, le lecteur pourra se référer à l'excellent site internet édité par Kolb *et al.* [Kolb 02], lequel contient de nombreuses références aux ouvrages essentiels.

A.2 Physiologie du système visuel humain

A.2.1 L'œil

L'œil est l'organe de la vision. Il permet à l'homme de capter la lumière pour ensuite l'analyser et interagir avec son environnement. Il a la forme d'une sphère d'environ 24 mm de diamètre. La figure A.1 présente une vue en coupe de l'œil droit suivant une section horizontale.

La paroi du globe oculaire est formée de trois tuniques :

- La tunique fibreuse, externe, composée de la *sclérotique* opaque en arrière et de la *cornée* transparente en avant. La sclérotique est ce qu'on appelle couramment le blanc de l'œil. Elle est traversée par un grand nombre de petits vaisseaux sanguins et par une ouverture à l'arrière où passent les fibres du *nerf optique*. La cornée est une lentille transparente dont le rôle est de capter et de focaliser la lumière sur le cristallin.
- La tunique uvéale, dite aussi uvée, est composée de trois éléments : l'*iris* en avant, le *corps ciliaire*, et la *choroïde* en arrière. L'iris est une membrane circulaire percée en son centre d'un orifice, la *pupille*. La contraction et la dilatation de l'iris permettent de modifier le diamètre de la pupille pour adapter la vision à la luminosité ambiante. Le corps ciliaire est un anneau de tissu musculaire qui produit une substance liquide appelée humeur aqueuse. Enfin, la choroïde est le tissu nourricier de l'œil : il apporte l'oxygène et les nutriments nécessaires au métabolisme des cellules.
- Enfin la tunique nerveuse, la plus interne, est composée essentiellement de la *rétine* qui permet la transformation du signal lumineux en signaux nerveux.

Ces tuniques renferment trois milieux transparents, dont l'association forment une lentille convergente :

- L'*humeur aqueuse* est un liquide transparent qui remplit l'espace entre la cornée et le cristallin. Ce liquide est continuellement renouvelé et permet, avec le corps vitré, le maintien de la pression oculaire.
- Le *cristallin* est la lentille principale de l'œil. C'est lui qui permet d'effectuer la mise au point et de focaliser le flux lumineux sur la rétine en modifiant sa géométrie lors de l'accommodation.
- Le *corps vitré* est une masse gélatineuse et transparente. Comme l'humeur aqueuse, son rôle est de maintenir la pression oculaire.

A.2.2 La rétine

La rétine est un tissu neuronal tapissant le fond de l'œil. C'est à ce niveau que s'effectue le premier traitement de l'information. Il consiste à traduire le message lumineux venant de l'extérieur en signaux nerveux utilisables et interprétables par les neurones des aires visuelles du cerveau. La rétine est composée de plusieurs couches de cellules neuronales comme illustré figure A.2.

La couche des cellules photosensibles, aussi appelées photo-récepteurs ou cellules sensorielles, est paradoxalement la couche la plus profonde de la rétine. Le signal lumineux doit en effet traverser l'ensemble des couches avant d'atteindre les photo-récepteurs. Cette couche est composée d'environ 130 millions de cellules réparties en deux types : les cônes et les bâtonnets, nommés ainsi en raison de leurs formes respectives.

La répartition des photo-récepteurs sur la rétine n'est pas uniforme (cf. figure A.3). Les cônes sont concentrés au centre de la rétine dans une zone d'environ 2 degrés de champ visuel appelée la *fovéa* (cf. figure A.1). La densité des cônes diminue rapidement lorsqu'on s'éloigne de la fovéa. Les bâtonnets sont environ 20 fois plus nombreux que les cônes mais sont peu représentés dans la fovéa

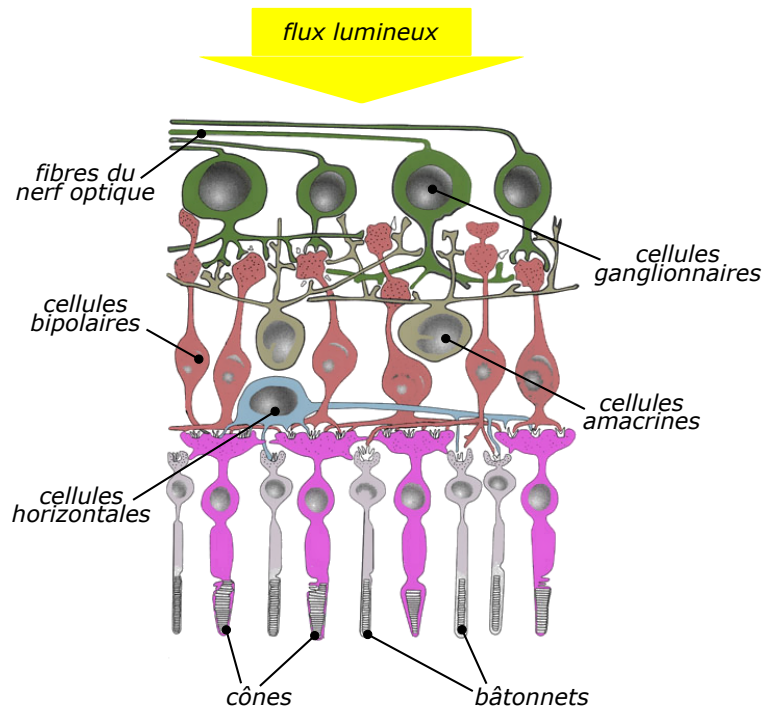


FIGURE A.2 – Structure de la rétine à proximité de la fovéa.

(ils sont même totalement absents d'une zone centrale de 0.7 degré appelée fovéola) et leur densité augmente à mesure que l'on s'en éloigne, avec un maximum pour une excentricité d'environ 20 degrés. Les signaux issus de plusieurs bâtonnets convergent vers une seule cellule ce qui augmente la sensibilité mais réduit la résolution. Au contraire, plusieurs cellules encodent l'information de chaque cône, ce qui confère à la fovéa une bien meilleure capacité en résolution spatiale que la zone périphérique.

Les cellules photo-réceptrices transmettent leur signal à un ensemble de couches de cellules (cf. figure A.2), véritables neurones rétiniens, qui assurent un premier traitement de l'information. L'information visuelle est ensuite acheminée via le nerf optique jusqu'au cortex visuel où sont effectuées les traitements haut-niveau de la vision.

A.3 Adaptation du système visuel humain à son environnement

Le système visuel humain a atteint au cours de l'évolution un très haut degré de complexité. Il est capable d'effectuer de nombreuses tâches que les technologies actuelles ne peuvent égaler. En premier lieu, le SVH est un système qui a su s'adapter à son environnement.

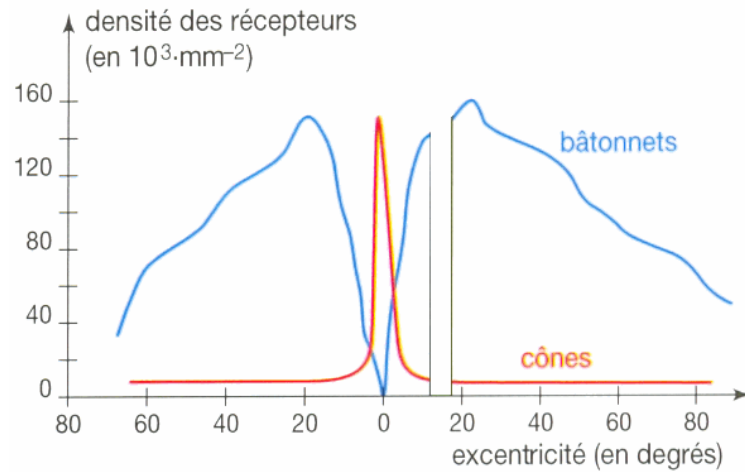


FIGURE A.3 – Densité des cellules photosensibles dans la rétine.

A.3.1 Sensibilité et adaptation à la lumière

A.3.1.1 Fonction d'efficacité lumineuse

La fonction d'efficacité lumineuse donne la sensibilité relative des photo-récepteurs aux différentes longueurs d'onde du spectre visible. On distingue deux fonctions d'efficacité lumineuse, une fonction photopique $V(\lambda)$ et une fonction scotopique $V'(\lambda)$, représentées figure A.4. En conditions photopiques (vision de jour, de 1 à 10^6 cd/m^2), la vision se fait principalement par l'intermédiaire des cônes. La sensibilité de l'œil est alors une combinaison des sensibilités spectrales des trois types de cônes, avec une nette domination des cônes L et M. La sensibilité maximale se situe à 555 nm, qui correspond à un rayonnement monochromatique vert. En conditions scotopiques (vision de nuit, de 10^{-6} à 10^{-2} cd/m^2), ce sont les bâtonnets qui sont principalement responsables de la vision, leur maximum de sensibilité se situe à environ 507 nm (bleu-vert). Enfin, en conditions mésopiques (de 10^{-2} à 1 cd/m^2) la fonction d'efficacité lumineuse est une combinaison des deux précédentes.

Ces fonctions d'efficacité permettent de convertir le flux radiant Φ d'une source électromagnétique (exprimé en watts) en flux lumineux F (exprimé en lumens) :

$$\begin{cases} F = K_m V(\lambda) \Phi(\lambda) & \text{en vision photopique} \\ F = K'_m V'(\lambda) \Phi(\lambda) & \text{en vision scotopique} \end{cases} \quad (\text{A.1})$$

En d'autres termes elles déterminent la proportion du flux radiant qui contribue à ses propriétés lumineuses. Plus généralement, cette fonction permet de passer d'une unité radiométrique à une unité photométrique. Les constantes K_m (vision photopique) et K'_m (vision scotopique) sont respectivement égales à 683 et 1700 lumens par watt.

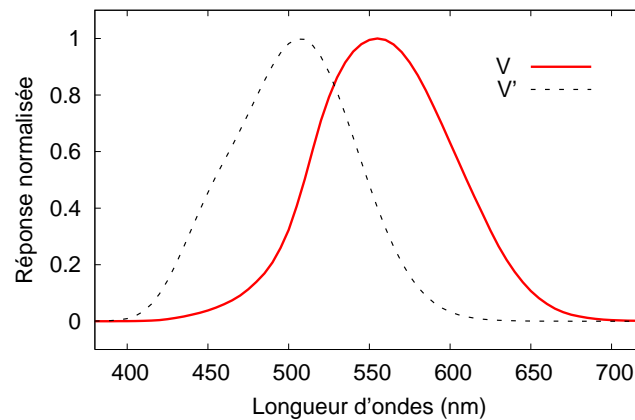


FIGURE A.4 – Fonctions d'efficacité lumineuse en conditions photopiques (V) et en conditions scotopiques (V') définie par le standard CIE 1931 [Wyszecki 82].

A.3.1.2 Adaptation à la lumière

Le système visuel humain peut s'adapter à une très large dynamique d'éclairement lumineux. Sa capacité d'adaptation s'étend de l'obscurité quasi-totale d'une nuit couverte, jusqu'à la luminosité extrême de la lumière du jour. En fait, de la vision scotopique à la vision photopique, l'œil peut s'adapter à une dynamique de luminance d'environ 12 ordres de magnitude (de 10^{-6} à 10^6 cd/m²). Pour atteindre une telle dynamique, plusieurs processus d'adaptation sont nécessaires. Le premier processus est la variation de la taille de la pupille (par contraction ou dilatation de l'iris, cf. section A.2.1). Le second, plus efficace, se situe au niveau des photo-récepteurs, via une modification de la concentration photo-chimique. A ce stade, deux phénomènes peuvent être distingués :

- Adaptation à la lumière : Le système visuel humain passe d'un environnement sombre à un environnement éclairé. L'adaptation se fait via la diminution de la concentration chimique dans les photo-récepteurs. Ce mécanisme est très rapide.
- Adaptation à l'obscurité : Le système visuel humain passe d'un environnement éclairé à un environnement sombre. Dans ce cas la concentration chimique, particulièrement celle de la rhodopsine, est augmentée. A l'inverse de l'adaptation à la lumière ce mécanisme est très lent et il peut se passer jusqu'à une heure avant que l'état stationnaire soit atteint.

Enfin, il existe également une adaptation neuronale, via une régulation du niveau du signal de sortie des photo-récepteurs. Ce mécanisme est très rapide mais significativement moins puissant que l'adaptation des photo-récepteurs.

A.3.1.3 Perception de l'intensité lumineuse

Indépendamment de ces mécanismes d'adaptation, la relation liant la luminance perçue et la luminance réelle n'est pas linéaire. Albert H. Munsell fut un des pionniers dans la description de ce

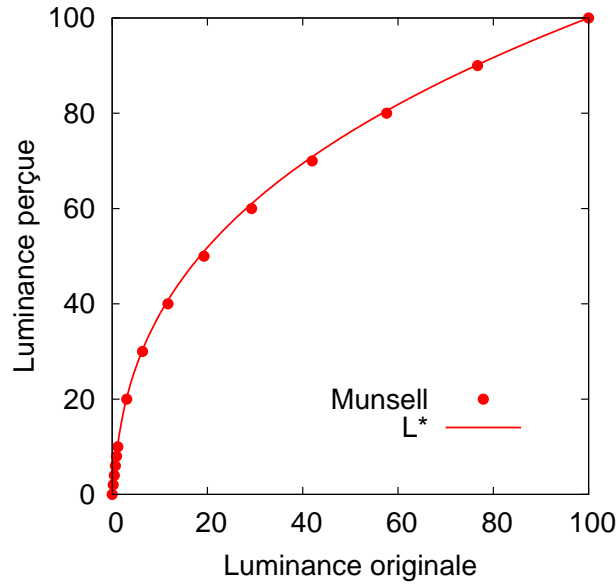


FIGURE A.5 – Relation entre la luminance originale et la luminance perçue : données expérimentales de Munsell [Munsell 15] (points) et luminance perçue L^* de l'espace de couleur CIE $L^*a^*b^*$.

phénomène [Munsell 15]. Grâce à des expérimentations intensives, il a construit la courbe figurant cette relation. De nombreux modèles existent, tous plus ou moins inspirés des expérimentations de Munsell. On peut citer notamment l'espace de couleur CIE $L^*a^*b^*$, dans lequel la luminance perçue, L^* est définie par :

$$L^* = \begin{cases} 116 \left(\frac{L}{L_n} \right)^{\frac{1}{3}} - 16 & \text{pour } \frac{L}{L_n} > 0.008856 \\ 903.3 \frac{L}{L_n} & \text{sinon} \end{cases} \quad (\text{A.2})$$

Le figure A.5 présente les données expérimentales de Munsell ainsi que le modèle de luminance perçue de l'espace de couleur CIE $L^*a^*b^*$.

A.3.2 Perception des couleurs

A.3.2.1 Aspect trichromatique de la vision des couleurs

Les études sur les cônes ont permis de dégager trois grandes populations qui se distinguent par leur sensibilité spectrale :

- les cônes S (pour *small*) ont une sensibilité maximale autour de 420 nm (bleu) ;
- les cônes M (pour *medium*) ont une sensibilité maximale autour de 531 nm (vert-jaune) ;
- enfin, les cônes L (pour *large*) ont leur maximum de sensibilité autour de 558 nm (jaune-rouge).

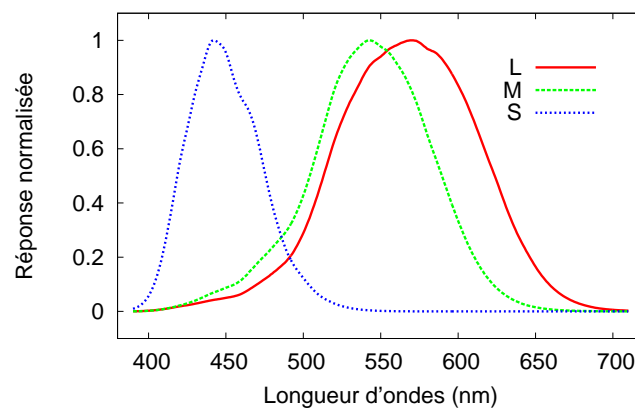


FIGURE A.6 – Réponse normalisée des trois types de photo-récepteurs, d'après Stockman et Sharpe [Stockman 99, Stockman 00].

La figure A.6 présente les réponses normalisées de ces trois types de cônes d'après les travaux de Stockman et Sharpe [Stockman 99, Stockman 00], eux-mêmes basés sur les travaux préliminaires de Stiles and Burch [Stiles 59].

Lorsque la répartition spectrale de la lumière est déséquilibrée au profit d'une ou plusieurs régions du spectre, la sensation lumineuse s'accompagne d'une sensation de couleur. Considérons un cône individuellement, un photon a une probabilité plus ou moins importante d'être absorbé par le cône, selon que sa longueur d'onde est plus ou moins proche du maximum de sensibilité spectrale de ce type de cellule photo-réceptrice. Cependant, tout photon absorbé déclenche un signal dans le cône, quelle que soit la longueur d'onde. On dit que la réponse du cône est univariante, i.e. elle ne peut varier qu'en amplitude bien que la lumière stimulante soit caractérisée par son amplitude mais aussi par sa longueur d'onde. Toute lumière provoque donc l'apparition de trois signaux au niveau des cônes, l'amplitude de ces signaux étant pondérée par la répartition spectrale de la lumière. Notre perception des couleurs dépend donc de cette combinaison de signaux, ce qui fait son aspect trichromatique.

A.3.2.2 Théorie des signaux antagonistes

Certains aspects de la perception des couleurs ne peuvent être expliqués par la seule théorie trichromatique. Goethe en 1810 [von Goethe 10] puis surtout Hering en 1878 [Hering 78] vont contribuer à l'établissement de la théorie des couleurs opposées. Cette théorie s'oppose à la théorie trichromatique en ce sens qu'elle prétend qu'il existe quatre couleurs fondamentales qui s'opposent deux par deux. Hering explique notamment que le jaune peut avoir une dominante rouge ou une dominante verte, mais en aucun cas une dominante bleue pour la bonne raison que ces couleurs s'opposent, ce

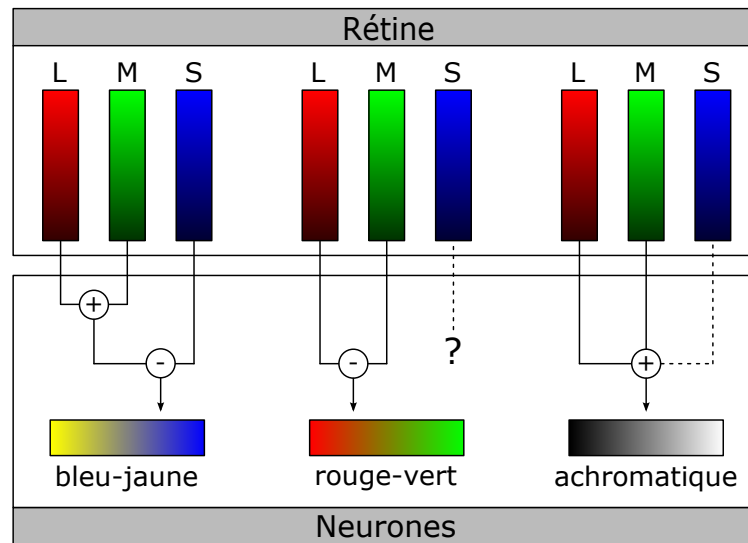


FIGURE A.7 – Schéma de principe des processus rétiens de vision des couleurs.

sont des couleurs antagonistes dont le mélange annule toute possibilité de couleurs autres que le gris ou le blanc. Hering en conclut qu'il n'y a pas trois couleurs fondamentales, mais quatre primaires correspondantes à des sensations psychologiques. C'est la théorie des couleurs opposées dont les 4 couleurs fondamentales sont le bleu, le vert, le jaune et le rouge.

Ainsi, les signaux issus des trois types de cônes (L, M, S) sont combinés au niveau des aires visuelles supérieures du cortex (cf. figure A.7). Trois axes sont ainsi définis :

- un axe chromatique bleu-jaune qui fournit une réponse antagoniste opposant les cônes S et la réponse additive issue des cônes M et L.
- un axe chromatique vert-rouge qui fournit une réponse antagoniste opposant les cônes L et les cônes M.
- axe achromatique qui fournit une réponse additive issue de la sollicitation des cônes L, M et S représentant une information sur le niveau d'éclairement indépendamment de la composition spectrale du rayonnement reçu.

Ces trois axes conduisent respectivement à deux composantes couleurs, nommées $Cr1$ et $Cr2$ et à une composante achromatique notée A . Différentes modélisations de la perception des couleurs par le SVH ont été proposées, majoritairement inspirées par les données biologiques et psychophysiques. La plupart du temps, ces modèles consistent en une transformation linéaire de la réponse des trois types de cônes :

$$\begin{pmatrix} A \\ Cr1 \\ Cr2 \end{pmatrix} = [T] \begin{pmatrix} L \\ M \\ S \end{pmatrix}$$

A.3.3 Mouvements de l'œil

Pour analyser notre environnement nos yeux sont continuellement en mouvement. La principale raison est que la fovéa, la zone haute résolution au centre de la rétine, ne couvre que 2 degrés de notre champ visuel. Par conséquent, le mouvement de nos yeux permet de recentrer les différentes zones d'intérêt de notre environnement afin de pouvoir les analyser de façon optimale. On distingue quatre types de mouvement oculaires.

A.3.3.1 Saccades

Les saccades sont des mouvements oculaires dont la vitesse est comprise entre 100 et 700 degrés par seconde. Ce type de mouvement permet de déplacer le regard d'une zone du champ visuel à une autre (un saut d'un point à un autre) afin d'inspecter ces zones par la partie la plus performante (en terme de résolution spatiale) de la rétine : la fovéa. Les saccades sont souvent considérées comme un mécanisme favorisant la sélection des informations visuelles pertinentes de notre champ visuel. La scrutation de notre monde visuel se fait donc par une série de sauts permettant le déplacement rapide de nos ressources sensorielles d'un point à un autre. Lorsqu'une saccade est effectuée en direction d'une cible particulière, la précision de la visée peut être soit très bonne soit mauvaise ; dans ce dernier cas, une seconde saccade ajuste le déplacement. Durant ces déplacements, notre pouvoir d'analyse est très faible signifiant que quasiment aucune information visuelle n'est traitée. Notons que le passage d'un point à un autre ne se fait pas forcément par le plus court chemin, c'est à dire la ligne droite. La trajectoire peut en effet être incurvée. En conditions normales, on compte généralement 2 à 3 saccades par secondes, elles sont séparées par des phases de fixations.

A.3.3.2 Fixations

Une phase de fixations se produit lorsque l'oeil fixe une zone du champ visuel. À première vue, l'œil a une position stationnaire d'où le terme de fixation. Pourtant et paradoxalement, les fixations sont considérées comme des mouvements oculaires. En effet, lors d'une phase de fixation, l'oeil est animé d'un mouvement résiduel. Ces légers mouvements, ou micro-saccades, permettent de décaler la zone examinée par la fovéa afin que cette dernière soit constamment excitée. Si l'œil était réellement stationnaire, la perception visuelle disparaîtrait progressivement. Des études [Watanabe 68, Kelly 79a] ont ainsi montré que l'œil perdait plus de 90% de sa sensibilité en vision complètement stabilisée (par compensation des micro-saccades).

A.3.3.3 Poursuite

Ce type de mouvement oculaire permet de suivre un objet en mouvement. Son rôle est important puisque dans ces conditions la vitesse relative de l'objet (la différence entre la vitesse de l'objet et celle de l'œil) est nulle. L'objet est alors stabilisé sur la rétine et peut être examiné par la fovéa avec un fort pouvoir de résolution. Comme dans le cas des fixations, l'objet n'est pas réellement immobile sur la rétine car malgré la poursuite l'œil se déplace en avant et en arrière de l'objet

suivi de manière à garder la fovéa continuellement excitée. La vitesse angulaire limite de poursuite visuelle est d'environ 30 deg/s [Yuyama 82, Daly 98].

A.3.3.4 Convergence

C'est un mouvement pour lequel les yeux se déplacent dans des directions horizontales opposées. Ce type de mouvement est utile pour acquérir des informations visuelles d'un objet proche ou éloigné de notre regard. Par exemple, lorsqu'un objet s'approche de notre regard, les yeux ajustent leurs positions en se rapprochant pour conserver une vision binoculaire, lorsque l'objet s'éloigne, les yeux s'écartent.

A.4 Limitations du système visuel humain

Malgré cette adaptation à son environnement, le SVH possède certaines limites intrinsèques qui influencent la qualité de l'image perçue. Dans le cadre de notre étude, nous allons nous intéresser plus particulièrement à la résolution spatiale et à la résolution temporelle, en effet ce sont les deux principaux facteurs qui permettent de percevoir une image nette et un mouvement fluide là où un signal discret spatio-temporellement est affiché.

A.4.1 Résolution spatiale et acuité visuelle

L'acuité visuelle peut être définie comme la capacité de résolution spatiale du système visuel humain, autrement dit la capacité de l'œil à discerner des détails fins. Plusieurs facteurs limitent cette acuité : les caractéristiques optiques de l'œil, la densité des cellules photo-réceptrices et la taille des champs récepteurs rétiniens.

A.4.1.1 Résolution optique

Le signal lumineux issu de la scène visuelle est acheminé vers la rétine via le système optique de l'œil. Les rayons lumineux traversent successivement la cornée, l'humeur aqueuse, le cristallin, et le corps vitré, avant de d'atteindre la rétine (cf. figure A.2). L'œil n'est pas un système optique parfait, les rayons lumineux sont légèrement déformés par la diffraction introduite par les différents milieux traversés. Ainsi, si l'on considère une source ponctuelle, son image ne sera pas ponctuelle mais distribuée sur la rétine (figure A.8a) selon un motif appelée la fonction d'étalement du point (en anglais *Point Spread Function*, PSF). Le système optique constitué par l'œil n'est donc pas stigmatique, on parle dans ce cas de stigmatisme approché.

La fonction d'étalement du point décrit le motif formé par une source ponctuelle sur la rétine (figure A.8b). Ce motif est appelé « disque d'Airy » ou « tâche d'Airy » du nom de l'astronome, Sir George B. Airy, qui le premier a réussi à en calculer le rayon exact. Ce rayon angulaire α (en radians) s'exprime en fonction de la longueur d'onde λ de la lumière et du diamètre d de la pupille selon la relation suivante :

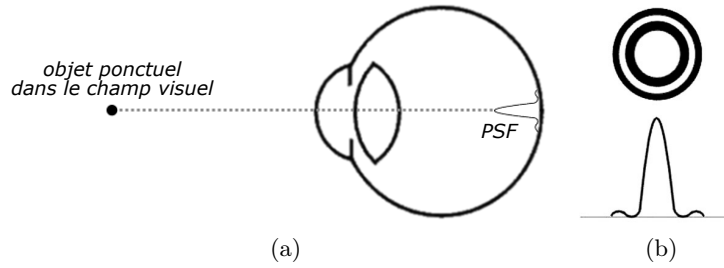


FIGURE A.8 – Fonction de d'étalement du point (PSF).

$$\alpha = 1.22 \frac{\lambda}{d}. \quad (\text{A.3})$$

A.4.1.2 Mosaïque rétinienne

Les dimensions et l'agencement des cellules photo-réceptrices sur la rétine limitent de façon intrinsèque l'acuité du système visuel humain. Ainsi, deux lignes verticales seront discriminées s'il existe une ligne de photo-récepteurs non stimulés entre les deux lignes de photo-récepteurs stimulés. Autrement dit, la résolution maximale de l'œil dépend directement de la taille et de la densité des détecteurs sur la rétine, c'est à dire des cônes dans la zone fovéale.

Lorsque l'on s'affranchit des distorsions optiques décrites dans la section précédente, la résolution spatiale mesurée est de l'ordre de 60 cycles par degré [Campbell 65]. Cependant, la densité des cellules photo-réceptrices conduit à une résolution théorique de la mosaïque rétinienne de l'ordre de 150 cycles par degré. La différence est due aux processus neuronaux qui entrent en jeu derrière la réception de signal effectuée par les cônes. En effet, les couches supérieures des cellules rétinienne possèdent elles aussi leur résolution propre, qui peut être modélisée par la notion de champ récepteur.

A.4.1.3 Cumul spatial et champs récepteurs

Le cumul spatial décrit la capacité de la rétine à additionner le flux lumineux reçu sur une certaine surface. La loi de Ricco décrit ce phénomène de la façon suivante : pour une surface inférieure à une certaine taille critique, le seuil de détection du flux lumineux est atteint quand l'énergie lumineuse totale atteint une certaine constante k . En d'autres termes, quand la luminance est divisée par deux, la surface du stimulus doit être doublée pour que celui-ci soit détecté. La loi de Ricco s'écrit de la façon suivante :

$$L \cdot A^n = k \quad (\text{A.4})$$

où L est la luminance du stimulus, A est la surface du stimulus et n décrit si la sommation spatiale est complète ($n = 1$) ou partielle ($0 < n < 1$). Le cumul spatial est incomplet lorsque la surface du stimulus dépasse une certaine surface critique. Ce phénomène est illustré figure A.9. Au delà de

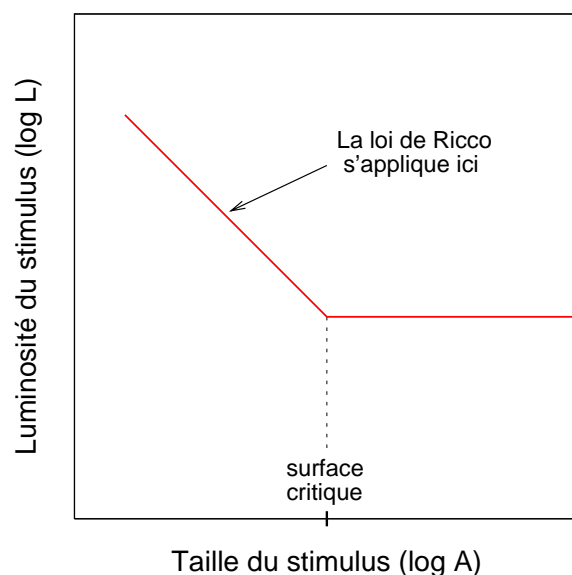


FIGURE A.9 – Illustration du phénomène de cumul spatial. La détection d'un stimulus dépend de sa luminance L et de sa surface A selon la loi de Ricco pour une surface inférieure à la surface critique.

la surface critique, l'augmentation de la taille du stimulus n'a plus d'influence sur la détection de celui-ci.

Les cellules photo-réceptrices transmettent leur signal à un ensemble de couches de cellules (cf. figure A.2), véritables neurones rétiniens, qui assurent un premier traitement de l'information. On trouve quatre types de cellules, les cellules horizontales, les cellules bipolaires, les cellules amacrines et les cellules ganglionnaires, ces dernières constituant l'ultime couche de cellules rétiniennes. Les interconnexions entre ces cellules sont à l'origine de la notion de champ récepteur, très importante pour la compréhension des mécanismes de perception du système visuel. Le champ récepteur d'un neurone est défini comme étant la zone pour laquelle la lumière influence la réponse du neurone. Ce concept n'est pas limité aux cellules de la rétine mais s'étend à tous les neurones des organes suivants du système visuel.

Le cumul spatial est dû à la convergence de plusieurs cellules photo-réceptrices sur un même neurone rétinien. L'excitation de plusieurs photo-récepteurs dans un champ récepteur donné ne produira qu'un seul signal. La taille des champs récepteurs augmente avec l'excentricité, ce qui explique pourquoi la surface critique du cumul spatial augmente elle aussi au fur et à mesure que l'on s'éloigne de la fovéa. On observe également une augmentation des champs récepteurs lorsque l'éclairement ambiant diminue. C'est donc la taille des champs récepteurs qui limite les capacités de résolution spatiale du système visuel humain.

A.4.1.4 Mesure de l'acuité visuelle

La mesure de l'acuité visuelle d'un observateur peut s'effectuer de différentes façons. On distingue généralement quatre types de tâches illustrées par la figure A.10 :

Reconnaissance de stimuli La tâche de l'observateur est de reconnaître un type de stimulus. On utilise généralement des tableaux de lettres, par exemple l'échelle optométrique de Monoyer en France (figure A.10a). L'acuité qui correspond à la taille, en minutes d'arc, des détails critiques du plus petit stimulus reconnu. Une acuité de 1 (ou 10/10) correspond donc à la reconnaissance de stimuli dont les détails font ont une dimension de 1 minute d'arc. L'acuité d'un observateur peut être supérieur à 10/10 si celui-ci parvient à lire des lettres dont les détails les plus fins ont une taille inférieure à 1 minute d'arc. On considère généralement que l'acuité d'un observateur moyen est égale à 1 (10/10).

Détection de stimuli Dans ce cas, on évalue la perception de l'absence ou de la présence d'un stimulus. L'acuité visuelle mesurée par ces tests correspond au diamètre apparent du point noir le plus petit, ou à la largeur de la ligne noire la plus fine, qui peuvent être distingués sur fond blanc (cf. figure A.10b, haut). D'autres stimuli standard existent, les optotypes, notamment l'anneau de Landolt (cf. figure A.10b, bas). La tâche de l'observateur consiste alors à détecter la brisure.

Résolution de stimuli L'acuité est alors exprimée comme le plus petit angle sous lequel les sujets peuvent différencier deux stimuli (deux points, une grille, ou encore un échiquier) comme illustré figure A.10c.

Localisation de stimuli Ces tests impliquent la détection d'une différence dans le positionnement (généralement l'alignement) des motifs d'un stimulus (cf. figure A.10d). L'acuité mesurée est alors l'angle minimum permettant de percevoir de la différence de position.

On parle généralement d'hyperacuité pour désigner les mesures d'acuité obtenues par détection, résolution ou localisation de stimuli. En effet, les valeurs obtenues sont généralement très inférieures à celles obtenues par reconnaissance de stimuli (acuité visuelle normale).

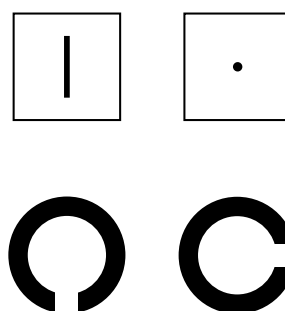
A.4.1.5 Facteurs affectant l'acuité visuelle

En dehors des deux principales limitations décrites ci-dessus, l'acuité visuelle dépend également de nombreux autres facteurs, parmi lesquelles on peut citer :

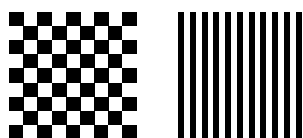
Les anomalies de l'œil réduisent l'acuité visuelle. La myopie et l'hypermétropie, en particulier, provoquent la formation de l'image respectivement devant et derrière la rétine.



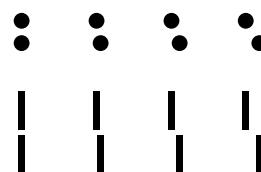
(a) Reconnaissance de stimuli.



(b) Détection de stimuli.



(c) Résolution de stimuli.



(d) Localisation de stimuli.

FIGURE A.10 – Exemples de stimuli utilisés pour la mesure de l'acuité visuelle, classés par type de tâche.

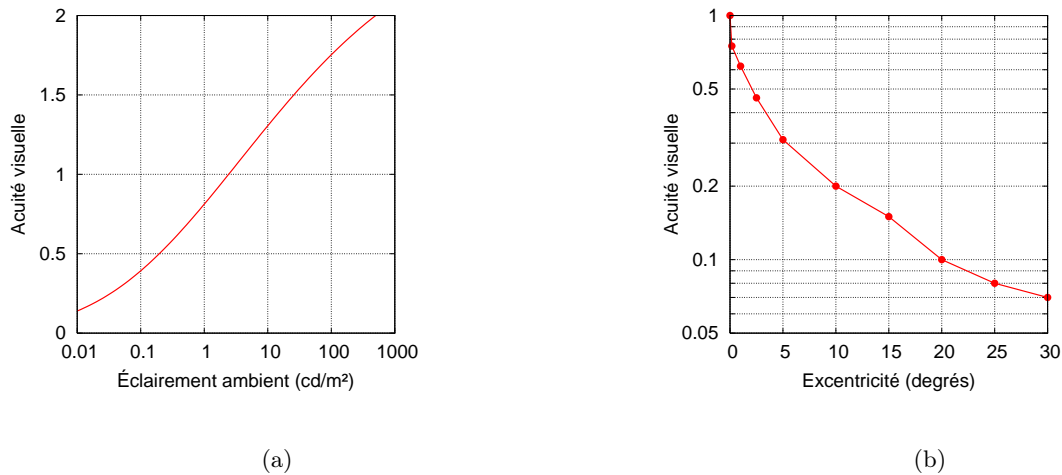


FIGURE A.11 – Influence de l'éclairement ambiant et de l'excentricité sur l'acuité visuelle. (a) Évolution de l'acuité visuelle en fonction de l'éclairement ambiant. L'acuité est mesurée en zone fovéale, par détection de stimuli, en utilisant l'anneau de Landolt (cf. figure A.10b bas) [Ikeda 80]. (b) Évolution de l'acuité visuelle en fonction de l'excentricité. L'acuité est mesurée sur différentes zones de la rétine, par reconnaissance de stimuli, en utilisant le tableau de lettres de Snellen (équivalent anglo-saxon du tableau de Monoyer) [Westheimer 72].

L'éclairement ambiant est un facteur important, particulièrement pour les tâches de reconnaissance de stimuli. La figure A.11a illustre ce phénomène pour une acuité visuelle mesurée par détection de stimuli. Dans une zone d'environ 4 ordre de magnitude (de 0.1 à 1000 cd/m^2), l'acuité visuelle varie linéairement en fonction du logarithme de l'éclairement ambiant.

La durée de présentation du stimulus joue un rôle important dans la mesure de l'acuité visuelle. Cependant, il n'y a pas de relation simple entre l'acuité et le temps d'exposition car cela dépend grandement du stimulus utilisé. L'acuité mesurée par la détection d'une ligne sera par exemple proportionnelle à la durée, mais dans le cas de la détection d'un point la quantité de lumière sera un facteur plus important que la durée.

La zone de la rétine impliquée influence fortement l'acuité visuelle. En raison de la forte densité de cônes dans la fovéa (cf. A.2.2), l'acuité visuelle est maximale au centre du champ visuel. En s'éloignant du centre de fixation, l'acuité décroît de façon exponentielle (cf. figure A.10b). Pour une excentricité de 10 minutes d'arc, on observe déjà une diminution de 25% de l'acuité. Il existe une forte corrélation entre la densité des cônes et l'acuité mesurée jusqu'à une excentricité d'environ 2 degrés. Au delà, l'acuité visuelle est inférieure à celle prédite par l'espacement des cônes, certainement limité par la taille des champs récepteurs impliqués dans le cumul spatial.

A.4.2 Résolution temporelle

Le flux lumineux issu de la scène visuelle est projetée sur la rétine de façon continue. Les informations visuelles sont ensuite intégrées par le SVH de manière à détecter les variations de luminance de la scène. Étant donné que la récupération et le traitement de ces informations ne sont pas instantanés, la réactivité de notre système visuel est limitée. Ainsi, le clignotement d'un stimulus sera perçu si la fréquence de ce clignotement est inférieure à une certaine valeur. Au delà de cette valeur appelée la fréquence de papillotement critique (en anglais *Critical Flicker Frequency*, CFF), le clignotement n'est plus perceptible et le stimulus apparaît constant. La cause principale de ce phénomène est la sommation temporelle.

A.4.2.1 Cumul temporel

La résolution temporelle du système visuel humain dépend du temps d'intégration des photo-récepteurs. En première approximation, on peut supposer que pour qu'un photo-récepteur détecte indépendamment deux flashes de lumière séparé d'un intervalle de temps Δt , sa période d'intégration devrait être strictement inférieure à $\Delta t/2$. La période d'intégration des photo-récepteurs peut monter jusqu'à 100 ms pour les bâtonnets alors qu'il n'est que de 10 ms pour les cônes. L'avantage d'un long temps d'intégration est que le seuil de détection peut être atteint sous un éclairement limité. En revanche, pour des niveaux d'éclairement plus importants un temps d'intégration plus court permet d'augmenter la résolution temporelle.

En réalité, la période d'intégration des photo-récepteurs n'est pas fixe, elle dépend du cumul temporel qui intervient au niveau des cellules rétinienne. Ce dernier évoque la propriété du système visuel humain à additionner les effets des quanta de lumière durant une certaine période de temps, appelée la durée critique. La loi de Bloch énonce que pour un stimulus d'une durée inférieure à cette durée critique, le seuil de détection est atteint lorsque l'énergie lumineuse totale atteint une certaine valeur constante k . En d'autres termes, le seuil est atteint lorsque le produit de la luminance L du stimulus et de sa durée T est égale à cette constante. Cette loi est analogue à la loi de Ricco sur la sommation spatiale.

La sommation temporelle cesse au delà de la durée critique, le seuil est alors seulement dépendant de la luminance du stimulus plutôt que du produit de sa luminance et de sa durée. La sommation temporelle est influencée par d'autres facteurs tels que l'éclairement ambiant et la taille du stimulus. La durée critique est plus importante sous un faible éclairement. En conditions scotopiques, la résolution temporelle du système visuel humain est donc plus faible. De même, lorsque la taille du stimulus est diminuée, la durée critique augmente indiquant une résolution temporelle plus faible.

A.4.2.2 Fréquence de papillotement critique

La fréquence de papillotement critique (CFF) est la fréquence de transition à partir de laquelle le papillotement (ou clignotement) d'une source intermittente n'est plus perçu. Au delà de cette fréquence, la source intermittente apparaît donc comme une source continue. Ce phénomène est une

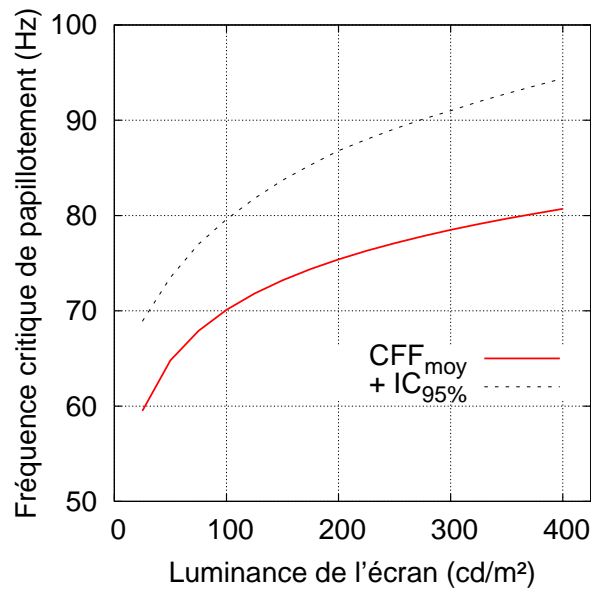


FIGURE A.12 – Fréquence critique de papillotement (CFF) d'un écran à tube cathodique en fonction de la luminance de l'écran. La courbe pleine représente la CFF moyenne, et la courbe en pointillés représente la partie haute de l'intervalle de confiance à 95% [ISO 91].

conséquence directe de la sommation temporelle. Un grand nombre de facteurs affectent donc notre perception du papillotement, notamment l'intensité et la taille du stimulus.

La loi de Ferry-Porter [Kolb 02] statue que la CFF est proportionnelle au logarithme de la luminance du stimulus clignotant :

$$CFF = a \cdot \log(L) + b \quad (\text{A.5})$$

En zone fovéale, cette relation reste vraie pour une large gamme de luminances (de 10^{-2} à 10^3 cd/m^2). Au fur et à mesure que l'intensité du stimulus augmente, nous sommes plus sensibles au papillotement. Dans des conditions d'éclairement normales, on admet généralement que la CFF se situe aux alentours de 50 Hz, mais elle peut atteindre 100 Hz pour des lumières intenses. La figure A.12 donne un exemple les valeurs de CFF mesurées dans des conditions similaires aux conditions d'utilisation d'un écran à tube cathodique en bureautique, d'après l'Organisation Internationale de Normalisation (ISO) [ISO 91]. On remarque par exemple qu'un écran CRT avec une luminance de 100 cd/m^2 doit avoir une fréquence de rafraîchissement de 70 Hz pour que le clignotement ne soit plus perçu par l'observateur moyen.

La loi de Ferry-Porter ne s'applique plus en dehors de la fovéa. En effet, si la CFF en zone fovéale ne dépend que de la luminance, en zone périphérique elle dépend de la proportion de bâtonnets et

de cônes excités par le stimulus car la résolution temporelle des bâtonnets est plus faible que celle des cônes. Pour de fortes luminances, ce sont les cônes qui déterminent la CFF et celle-ci diminue fortement avec l'excentricité. Pour de faibles luminances, ce sont les bâtonnets qui déterminent la valeur de la CFF et celle-ci est plus élevée qu'en zone fovéale où les bâtonnets sont absents.

Par conséquent, la CFF varie fortement en fonction de la taille du stimulus. En effet, même s'il est centré sur la fovéa, un stimulus de grande taille va exciter les zones excentrées de la rétine. Par conséquent, pour de faibles luminances un grand stimulus excitera les bâtonnets situés en zone péri-fovéale, et la CFF sera plus élevée que pour un petit stimulus. De plus, plus la taille du stimulus augmente plus le nombre de cellules photo-réceptrices excitées est important. En conséquence, la résolution temporelle augmente avec la taille du stimulus. En conditions photopiques, et pour une luminance donnée, on considère généralement que la CFF est proportionnelle au logarithme de la taille du stimulus.

A.5 Sensibilité au contraste et aux textures

A.5.1 Loi Weber-Fechner

Le système visuel humain peut percevoir de très faibles différences de luminance. La mesure de l'incrément lumineux nécessaire à la perception d'une différence peut se faire via une expérience simple. Considérons un stimulus constitué d'un fond uniforme de luminance L sur lequel est présenté un patch de luminance $L + \Delta L$ (cf. figure A.13a). Le contraste de ce stimulus est défini comme le rapport entre l'incrément de luminance et la luminance de fond :

$$C = \frac{\Delta L}{L} \quad (\text{A.6})$$

Pour définir la sensibilité au contraste du système visuel humain, l'expérience consiste à déterminer l'incrément de luminance ΔL minimum nécessaire pour percevoir le patch. La valeur de cet incrément dépend de la luminance de fond L . Selon la loi Weber-Fechner, pour une luminance de fond comprise entre 0.1 et 1000 cd/m², l'incrément de luminance nécessaire à la perception du patch est proportionnel à la luminance de fond :

$$\Delta L = kL \quad (\text{A.7})$$

Autrement dit, le contraste minimum perceptible par l'œil humain est constant :

$$C = \frac{\Delta L}{L} = k \quad (\text{A.8})$$

Sa valeur est comprise entre 1% et 3%.

La relation entre le contraste seuil et la luminance de fond est donnée figure A.13b. On remarque la région plate correspondant à la loi de Weber-Fechner. Pour des luminances de fond plus faibles et plus fortes, un contraste plus élevé est nécessaire pour détecter le patch. Par ailleurs, la sensibilité au contraste dépend fortement de la luminance d'adaptation de l'œil. Le contraste nécessaire pour la

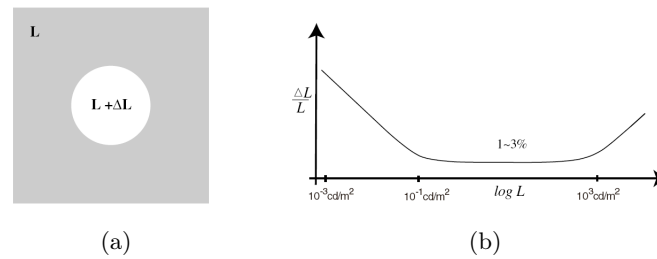


FIGURE A.13 – La sensibilité au contraste et la loi de Weber-Fechner.

détection du patch sera minimum si l'œil est adapté à la luminance de fond L , pour des luminances d'adaptation différentes – supérieures ou inférieures – un contraste plus élevé est nécessaire.

A.5.2 Fonctions de sensibilité aux contrastes (CSF)

On vient de le voir, la sensibilité au contraste pour des stimuli simples et uniformes peut être décrite de façon simple. Cependant, cette description n'est pas adaptable à notre environnement visuel. En effet, les images perçues par nos yeux sont constituées d'une multitude de détails, de formes, de contours, et de mouvements. Il est donc nécessaire d'étudier la sensibilité au contraste en fonction de cette activité spatiale et temporelle.

La figure A.14 permet de se rendre compte de façon visuelle que la sensibilité au contraste dépend de la fréquence spatiale. La luminance des pixels varie de façon sinusoïdale dans la direction horizontale. La fréquence spatiale de cette sinusoïde augmente de façon exponentielle de la gauche vers la droite et son amplitude (son contraste) diminue de façon exponentielle du bas vers le haut. On remarque immédiatement que le contraste pour lequel le stimulus n'est plus visible varie en fonction de la fréquence spatiale. Le système visuel humain est plus sensible aux moyennes fréquences et sa sensibilité décroît pour les basses et hautes fréquences spatiales. On désigne par fonction de sensibilité au contraste (*Contrast Sensitivity Function*, CSF), les fonctions qui donnent la sensibilité à un stimulus en fonction des caractéristiques de ce dernier.

La sensibilité est définie comme l'inverse du contraste seuil pour lequel un stimulus devient imperceptible. Elle est mesurée via des expériences psychophysiques. Pour des raisons pratiques et physiologiques, on utilise généralement des motifs sinusoïdaux, permettant une analyse harmonique de la réponse du système visuel humain. Par exemple, le motif de Gabor (cf. figure A.15) est très utilisé. Ce stimulus consiste en une fonction sinusoïdale apodisée, dans les deux directions spatiales, par une fonction gaussienne. Ainsi la perception de l'observateur n'est pas perturbée par les hautes fréquences qui seraient introduites dans le cas d'une transition nette entre le motif sinusoïdal et le fond uniforme. Pour une orientation horizontale, l'expression mathématique du motif de Gabor est la suivante :

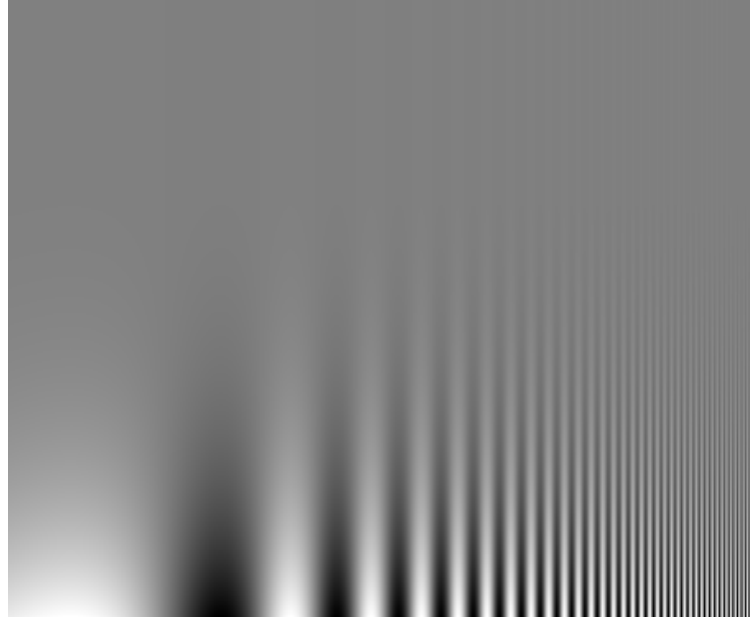


FIGURE A.14 – Diagramme de Campbell et Robson.

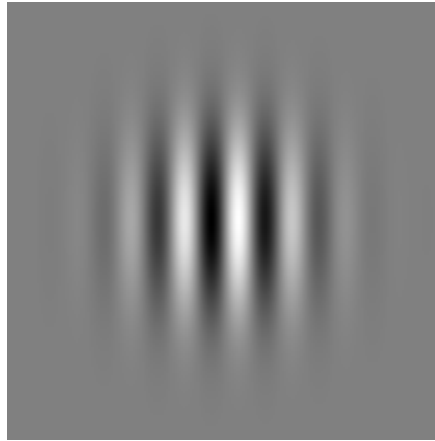


FIGURE A.15 – Motif de Gabor utilisé pour mesurer la sensibilité aux contrastes.

$$\text{Gabor}(x, y) = A \left[\sin(2\pi f_s x) \cdot \exp \left(-\frac{x^2 + y^2}{0.36 \cdot l^2} \right) \right] + L_{moy} \quad (\text{A.9})$$

où A est l'amplitude, f_s la fréquence spatiale, l la largeur à mi-hauteur de la fenêtre gaussienne, et L_{moy} la luminance moyenne du stimulus. On a $L_{moy} \geq A/2$. Le contraste du motif de Gabor s'exprime de la façon suivante (contraste de Michelson) :

$$C = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} = \frac{A}{2 \cdot L_{moy}} \quad (\text{A.10})$$

A.5.2.1 Fonction de sensibilité au contraste spatial (SCSF)

On désigne par fonction de sensibilité au contraste spatial (ou CSF spatiale), la fonction qui donne la sensibilité (inverse du contraste seuil) en fonction de la fréquence spatiale du stimulus. Un exemple de CSF spatiale est donné figure A.16 pour un stimulus achromatique (courbe noire) [Daly 93]. On remarque la forme passe-bande de la fonction de sensibilité au contraste. La sensibilité (l'inverse du contraste seuil) du système visuel humain est maximale (contraste seuil de l'ordre de 0.5%) pour une fréquence spatiale d'environ 3-4 cycles par degré. Elle diminue rapidement pour les hautes fréquences et légèrement pour les basses fréquences. La fréquence de coupure de la CSF se situe aux alentours de 50 cycles par degré.

La figure A.16 donne également les CSF obtenues avec des stimulus chromatiques (courbes rouge et bleue) [Le Callet 01]. Les courbes de sensibilité sont de forme passe-bas, et les valeurs de sensibilité sont inférieures à celles obtenues pour des stimuli achromatiques. De plus, on remarque que le système visuel humain est moins sensible à la composante Cr_2 (bleu-jaune) qu'à la composante Cr_1 (rouge-vert).

De nombreux facteurs influencent le résultats de ces expériences :

- La définition du contraste n'est pas unique et la sensibilité sera différente selon la définition choisie.
- La taille du stimulus et le nombre de cycles affichés modifient également la sensibilité, ces paramètres doivent donc être choisis avec précautions.
- L'éclairement ambiant est également un paramètre important, il a été montré que la sensibilité était maximale pour un éclairage proche de la luminance moyenne du stimulus.
- La forme de la CSF dépend également de l'orientation du stimulus. Des stimuli orientés horizontalement et verticalement donnent des résultats assez similaires, mais le système visuel humain est moins sensible aux orientations obliques.

A.5.2.2 Fonction de sensibilité au contraste spatio-temporel

La sensibilité au contraste peut également être mesurée le long de l'axe temporel. Dans ce cas, c'est la fréquence temporelle du stimulus qui varie. Il y a différentes façons de faire varier la fréquence temporelle. Les premières études sur ce sujet [Kelly 69, Robson 66] ont introduit une variation temporelle des motif sinusoïdaux par clignotement déphasé. Autrement dit, le stimulus spatial est modulé par une fonction sinusoïdale du temps. Si on reprend l'exemple du motif de Gabor donné par l'équation A.9, on obtient donc :

$$\text{Gabor}(x, y, t) = \text{Gabor}(x, y) \cdot \cos(2\pi f_t t) \quad (\text{A.11})$$

où f_t est la fréquence temporelle du clignotement.

Ce clignotement déphasé a été utilisé pour comprendre la sensibilité du système visuel humain aux variations rapides de luminance telles que celles rencontrées sur les téléviseurs ou sur les lampes

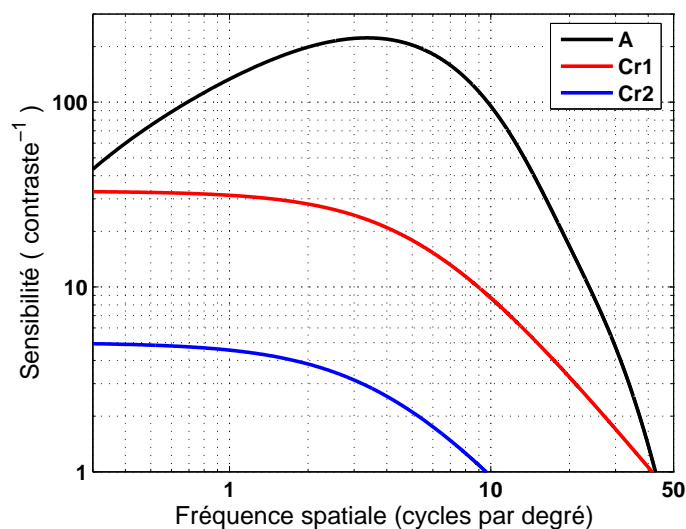


FIGURE A.16 – Fonction de sensibilité aux contrastes spatiaux pour la composante achromatique [Daly 93], et pour les deux composantes chromatiques [Le Callet 01].

fluorescentes. La figure A.17 présente des CSF spatiales pour plusieurs fréquences temporelles fixes (à gauche) et des CSF temporelles pour plusieurs fréquences spatiales fixes [Robson 66]. La similarité entre les deux figures est évidente. Il y a une baisse de la sensibilité pour les faibles fréquences spatiales uniquement pour les stimuli de faibles fréquences temporelles et vice-versa (cf. la courbe 1 Hz sur la figure de gauche et la courbe 0.05 cpd sur la figure de droite). Au fur et à mesure que la fréquence temporelle du stimulus est augmentée, la CSF spatiale varie d'une forme passe-bande vers une forme passe-bas. De même, à mesure que la fréquence spatiale augmente la CSF temporelle évolue d'une forme passe-bande vers une forme passe-bas. En fait, la forme de la CSF spatiale (respectivement temporelle) ne varie pas pour les hautes fréquences temporelles (respectivement spatiales). En fait, il a été montré [Kelly 69, Burbeck 80] que les CSF spatiales et temporelles sont séparables pour de hautes fréquences temporelles et spatiales (respectivement au dessus de 10 Hz et 10 cpd). Dans ce domaine de fréquences, la CSF spatio-temporelle est simplement le produit de la CSF spatiale et de la CSF temporelle. Pour des fréquences inférieures à 10 Hz et 10 cpd, des interactions dues au comportement temporel des photo-récepteurs et aux inhibitions latérales des champs récepteurs ne permettent pas de mesurer la sensibilité temporelle du système visuel indépendamment de la sensibilité spatiale et vice-versa.

A.6 Perception du mouvement

A.6.1 Généralités

La perception du mouvement est le processus qui permet au système visuel humain de déduire la vitesse et la direction des éléments présents dans la scène visuelle. Malgré sa simplicité apparente,

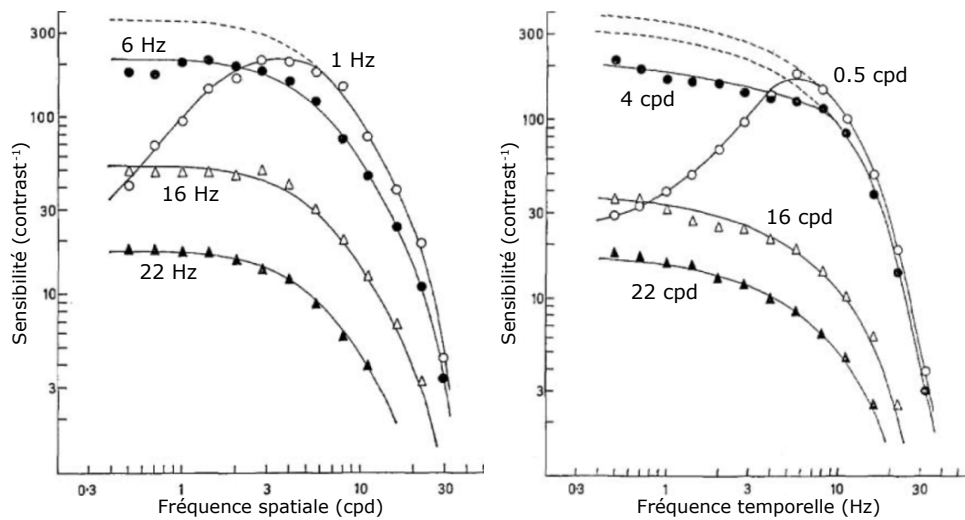


FIGURE A.17 – CSF spatiales pour des stimuli de différentes fréquences temporelles (droite), CSF temporelles pour des stimuli de différentes fréquences spatiales (gauche) [Robson 66].

ce traitement reste un problème difficile à modéliser et extraordinairement difficile à expliquer en termes de traitement neuronal. On distingue deux types de mouvements, qui diffèrent par la façon dont ils sont perçus par le système visuel humain.

Le mouvement de premier ordre est le mouvement d'un objet dont la luminance est différente de celle de l'environnement dans lequel il se déplace. Ce type de mouvement est détecté par des capteurs de mouvement relativement simple dont le rôle consiste à détecter un changement de luminance sur un point de la rétine et à le corrélérer avec un changement de luminance sur un point voisin de la rétine après un certain délai. Ces capteurs sont appelés détecteurs de Reichardt ou capteurs mouvement-énergie [Adelson 85, van Santen 85], ils fonctionnent par corrélation spatio-temporelle et sont un modèle plausible pour décrire la détection du mouvement par le système visuel. Les détecteurs de mouvement du premier ordre souffrent du « problème d'ouverture », en effet leur fonctionnement ne leur permet que de détecter des mouvements perpendiculaires à l'orientation des contours en déplacement. Des processus de plus haut niveau sont ensuite nécessaires pour lever l'ambiguïté sur la direction réelle du mouvement global.

Les mouvements de second ordre regroupent les mouvements définis par le contraste du stimulus, sa texture, son clignotement, ou tout autre caractéristique qui ne modifie pas l'énergie de son spectre de Fourier [Ledgeway 95, Hutchinson 06]. Il a été montré que les mouvements de premier et de second ordre sont traités séparément et acheminés par des voies différentes jusqu'au cortex visuel. Les mécanismes de second ordre ont une résolution temporelle plus faible et ont un comportement passe-bas pour les fréquences spatiales auxquelles ils répondent.

A.6.2 Cinéma et télévision : l'illusion du mouvement

Sur un écran cinéma comme sur une télévision, on présente à l'observateur une série d'images stationnaires. Pourtant, l'observateur perçoit un mouvement continu. Cette illusion du mouvement

repose sur deux principes visuels distincts, bien que souvent confondus : la persistance rétinienne et le phénomène bêta.

A.6.2.1 Persistance rétinienne

Par persistance rétinienne, on entend simplement l'incapacité de la rétine à signaler des changements rapides d'intensité lumineuse. Cette propriété est directement liée à la résolution temporelle de l'œil et au temps d'intégration des cellules photo-réceptrices (cf. section A.4.2.1). On a longtemps cru que le cinéma se servait uniquement de ce phénomène pour créer l'illusion du mouvement. Ce « mythe de la perception rétinienne » a encore la peau dure, et on trouve encore aujourd'hui de nombreuses publications, principalement dans le domaine du cinéma, qui font référence à la persistance rétinienne comme cause principale de l'illusion du mouvement, et ce malgré les efforts répétés des spécialistes de la vision pour tuer le mythe [Anderson 78, Anderson 93]

En réalité la persistance rétinienne, ou plutôt la sommation temporelle qui intervient au niveau rétinien, est seulement responsable de la disparition du clignotement des images affichées sur un écran de cinéma ou sur un téléviseur, à partir d'une certaine fréquence d'affichage. Cette limitation de l'œil a été expliquée en détail précédemment (cf. section A.4.2.2).

A.6.2.2 Le phénomène bêta

La perception du mouvement au cinéma, alors qu'on ne présente à l'observateur qu'une succession d'images fixes, est principalement due à une illusion perceptuelle appelée le phénomène bêta. Ce phénomène a été décrit la première fois en 1912, par Max Wertheimer l'un des fondateurs de la psychologie Gestalt ou psychologie de la forme. Cette illusion permet de tromper le cerveau. L'expérience classique pour démontrer ce phénomène consiste à présenter à un observateur deux images successives. La première image représente une balle sur la partie gauche de l'écran, la seconde représente une balle sur la partie droite de l'écran. Si on interroge l'observateur sur ce qu'il a vu, il répondra qu'il a vu une balle se déplacer de gauche à droite, alors qu'en fait ce n'est pas ce qu'il a vu.

La figure A.18 illustre le phénomène bêta. On affiche successivement les deux images décrites précédemment avec une fréquence que l'on augmente progressivement. Si l'intervalle de temps entre les deux images est trop long, l'observateur percevra chaque balle l'une après l'autre en succession, le mouvement n'est pas perçu. Si on augmente la fréquence d'affichage, l'observateur percevra un mouvement de va-et-vient de la balle, c'est le phénomène bêta. Si on affiche les images avec une fréquence très élevée, la fréquence critique de clignotement (cf. section A.4.2.2) sera dépassée pour chacune des balles et celles-ci apparaîtront simultanément.

Ici encore, de nombreuses confusions subsistent, notamment entre le phénomène bêta et un autre phénomène également découvert par Max Wertheimer : le phénomène phi [Steinman 00]. Ce phénomène est également décrit figure A.18. Il intervient entre le mouvement bêta et la simultanéité. Le phénomène phi est appelé le « mouvement pur ». Il n'a pas de forme, et prend la couleur du fond. Ici ce n'est pas la balle que l'on perçoit en mouvement mais un objet informe qui se déplace

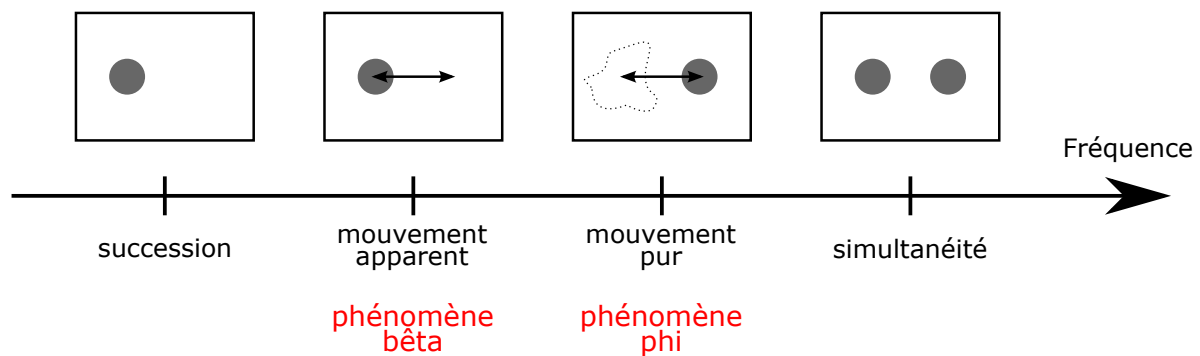


FIGURE A.18 – Illustration du phénomène bêta créant l’illusion du mouvement à partir de deux images fixes.

de gauche à droite et qui masque successivement chacun des deux objets. Contrairement à ce qu’on peut lire dans certains ouvrages [Gregory 00], le phénomène phi et le phénomène bêta sont bien distincts, et seul le phénomène bêta est responsable de l’illusion du mouvement au cinéma.

A.7 Conclusion

Dans cette annexe, les propriétés utiles et les limitations importantes du système visuel humain ont été décrites. Dans le cadre de ce travail de thèse, on peut particulièrement retenir les points suivants :

- La réponse du système visuel humain à la luminance, et aux variations de luminance, n’est pas linéaire.
- La mobilité de l’œil permet au système visuel humain de réaliser un mouvement de poursuite pour stabiliser un objet en mouvement au centre de la rétine.
- Le système visuel humain possède une résolution spatiale et une résolution temporelle limitées, qui dépendent de nombreux paramètres.
- La sensibilité au contraste du système visuel humain dépend de la fréquence spatiale et de la fréquence temporelle du stimulus. La sensibilité temporelle est de plus différente selon que la variation temporelle du stimulus est induite par un mouvement ou par clignotement.

Bibliographie

- [Abramson 03] Albert Abramson. The history of television, 1972 to 2000. McFarland & Company, 2003. ISBN 0786412208.
- [Adam 07] Pierre Adam, Pascal Bertolino & Fritz Lebowsky. *Mathematical modeling of the LCD response time*. Journal of the Society for Information Display, vol. 15, no. 8, pages 571–577, August 2007.
- [Adam 08] Pierre Adam. *Améliorations d’artefacts sur panneaux LCD*. PhD thesis, Institut National Polytechnique de Grenoble, 2008.
- [Adelson 85] E. H. Adelson & J. R. Bergen. *Spatiotemporal energy models for the perception of motion*. Journal of the Optical Society of America A, vol. 2, no. 2, pages 284–299, 1985.
- [Ahumada 98] Albert J. Ahumada, Bettina L. Beard & Robert Eriksson. *Spatiotemporal discrimination model predicts temporal masking functions*. In Proceedings of the SPIE Human Vision and Electronic Imaging III, volume 3299, pages 120–127, 1998.
- [Anderson 78] Joseph Anderson & Barbara Fisher. *The myth of persistence of vision*. Journal of the University Film Association, vol. XXX, no. 4, pages 3–8, 1978.
- [Anderson 93] Joseph Anderson & Barbara Anderson. *The myth of persistence of vision revisited*. Journal of Film and Video, vol. 45, no. 1, pages 3–12, 1993.
- [Aoki 96] Hotoshi Aoki. *Dynamic characterization of a-Si TFT-LCD pixels*. IEEE Transactions on Electron Devices, vol. 43, no. 1, pages 31–39, Janvier 1996.
- [Barten 99] Peter G. J. Barten. Contrast sensitivity of the human eye and its effects on image quality. SPIE Optical Engineering Press, 1999.
- [Becker 07] Michael E. Becker. *Evaluation of transition times from filtered luminance data*. White paper, Display-Metrology & Systems, Octobre 2007.
- [Becker 08a] Michael E. Becker. *LCD response time evaluation in the presence of backlight modulations*. In SID Symposium Digest of Technical Papers, volume 39, pages 24–27, 2008.
- [Becker 08b] Michael E. Becker. *Motion-blur evaluation : A comparison of approaches*. Journal of the Society for Information Display, vol. 16, no. 10, pages 989–1000, Octobre 2008.

- [Bellers 07] E. B. Bellers, J. G. Janssen & M. Penners. *Efficient frame-rate conversion for high frame rate LCD panels*. In Proceedings of the Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM2007, numéro 395, 2007.
- [Bex 07] Peter J. Bex & Keith Langley. *The perception of suprathreshold contrast and fast adaptive filtering*. Journal of Vision, vol. 7, no. 12, pages 1–23, September 2007.
- [Blin 06] Jean-Louis Blin. *New quality evaluation method suited to multimedia context : SAMVIQ*. In Proceedings of the Second International Workshop on Video Processing and Quality Metrics, VPQM'06, Scottsdale, January 2006.
- [Bowker 83] D. O. Bowker. *Suprathreshold spatiotemporal response characteristics of the human visual system*. Journal of the Optical Society of America A, vol. 73, no. 4, pages 436–440, April 1983.
- [Brainard 97] David H. Brainard. *The Psychophysics Toolbox*. Spatial Vision, vol. 10, no. 4, pages 433–436, 1997.
- [Bryngdahl 66] Olof Bryngdahl. *Characteristics of the visual system : psychophysical measurements of the response to spatial sine-wave stimuli in the photopic region*. Journal of the Optical Society of America, vol. 56, no. 6, pages 811–821, June 1966.
- [Burbeck 80] Christina A. Burbeck & D. H. Kelly. *Spatiotemporal characteristics of visual mechanisms : excitatory-inhibitory model*. Journal of the Optical Society of America A, vol. 70, no. 9, pages 1121–1126, September 1980.
- [Campbell 65] F.W. Campbell & D. G. Green. *Optical and retinal factors affecting visual resolution*. Journal of Physiology, vol. 181, pages 576–593, 1965.
- [Cannon 85] Mark W. Jr. Cannon. *Perceived contrast in the fovea and periphery*. Journal of the Optical Society of America A, vol. 2, no. 10, pages 1760–1768, October 1985.
- [Carrein 03] Geert Carrein. *Characteristics of CRT and LCD displays*. White paper, Barco, 2003.
- [Daly 93] Scott Daly. *The visible differences predictor : An algorithm for the assessment of image fidelity*, chapitre 14, pages 179–206. MIT press, 1993.
- [Daly 98] Scott Daly. *Engineering observations from spatiovelocity and spatiotemporal visual models*. In Proceedings of the SPIE Conf. Human Vision and Electronic Imaging III, volume 3299, pages 180–191. IS&T/SPIE, 1998.
- [DePalma 62] J. J. DePalma & E. M. Lowry. *Sine-wave response of the visual system. II. Sine-wave and square-wave contrast sensitivity*. Journal of the Optical Society of America A, vol. 52, no. 3, pages 328–335, March 1962.
- [Display Search 05] Display Search. *Advanced TV performance report*. Display Search and Westar Display Technologies, Août 2005.

- [EBU 03] EBU. *SAMVIQ - Subjective assessment methodology for video quality*. Rapport technique, European Broadcasting Union, 2003.
- [EBU 04] EBU. *Maximizing the quality of SDTV in the flat-panel environment*. EBU technical review, European Broadcasting Union - Broadcast Technology Management Committee, April 2004.
- [Farnsworth 37] Philo T. Farnsworth. *Image dissector*. United States Patent Office, July 1937. Patent number : 2087683.
- [Feng 06] Xiao-Fan Feng. *LCD motion-blur analysis, perception, and reduction using synchronized backlight flashing*. In Proceedings of the SPIE Conf. Human Vision and Electronic Imaging XI, volume 6057. Electronic Imaging 2006, Janvier 2006.
- [Feng 08] Xiao-Fan Feng, Hao Pan & Scott Daly. *Comparisons of motion-blur assessment strategies for newly emergent LCD and backlight driving technologies*. Journal of the Society for Information Display, vol. 16, no. 10, pages 981–988, 2008.
- [Fisekovic 01] N. Fisekovic. *Improved motion picture quality of AM-LCDs using scanning backlight*. In International Displays Workshops 2001, pages 1637–1640, 2001.
- [Fiser 03] József Fiser, Peter J. Bex & Walter Makous. *Contrast conservation in human vision*. Vision Research, vol. 43, pages 2637–2648, 2003.
- [Friedel 22] Georges Friedel. *Les états mésomorphes de la matière*. Annales de Physique, vol. 18, 1922.
- [Fujine 07] Toshiyuki Fujine, Yuhji Kikuchi, Michiyuki Sugino & Yasuhiro Yoshida. *Real-life in-home viewing conditions for flat panel displays and statistical characteristics of broadcast video signal*. Japanese Journal of Applied Physics, vol. 46, no. 3B, pages 1358–1362, Mars 2007.
- [Fujio 82] Takashi Fujio. *Future broadcasting and high definition television*. NHK technical monograph, NHK Technical Monograph, June 1982.
- [Georgeson 75] M. A. Georgeson & G. D. Sullivan. *Contrast constancy : deblurring in human vision by spatial frequency channels*. The Journal of Physiology, vol. 252, no. 3, pages 627–656, November 1975.
- [Georgeson 87] Mark A. Georgeson. *Temporal properties of spatial contrast vision*. Vision Research, vol. 27, pages 765–780, 1987.
- [Gregory 00] Richard L. Gregory, Micheline Mattheeuws-Hambrouck & Georges Thinès. *L'œil et le cerveau : la psychologie de la vision*. DeBoeck Université, 2000.
- [Gu 06] Mingxia Gu. *The World of Liquid Crystal Displays*. <http://www.personal.kent.edu/~mgu/LCD/index.htm>, 2006. Liquid Crystal Institute, Kent State University.
- [Haiyan 07] H. Haiyan, Leo Jan Velthoven, E. B. Bellers & J. G. Janssen. *Analysis and implementation of motion compensated inverse filtering for reducing motion blur on LCD panel*. In Proceedings of the International Conference on Consumer Electronics, ICCE 2007, 2007.

- [Hennessy 05] Brian Hennessy. The emergence of broadcasting in Britain. Southerleigh, 2005. ISBN 0955140803, 9780955140808.
- [Hering 78] Ewald Hering. Outlines of a theory of the light sense. 1878.
- [Hutchinson 06] Claire V. Hutchinson & Timothy Ledgeway. *Sensitivity to spatial and temporal modulations of first-order and second-order motion*. Vision Research, vol. 46, pages 324–335, 2006.
- [Ikeda 80] Koichi Ikeda, Koji Noda & Shoichiro Yamaguchi. *A relation between adaptation luminance and visual acuity for the Landolt ring under the uniform background*. Journal of light and visual environment, vol. 4, no. 2, pages 22–31, 1980.
- [Ishihara 67] Shinobu Ishihara. *Tests for Colour-Blindness*. Kanehara Shuppen Company, Ltd., Tokyo, Japan, 1967.
- [ISO 91] ISO. *Visual Display Terminals (VDTs) used for office tasks - Ergonomic requirements - Part 3 : Visual displays*. 9241 3, International Organization for Standardization, 1991.
- [ISO 01] ISO. *Ergonomic requirements for work with visual displays based on flat panels – Part 2 : Ergonomic requirements for flat panel displays*. 13406 2, International Organization for Standardization, 2001.
- [ISO 08] ISO. *Ergonomics of human-system interaction – Part 305 : Optical laboratory test methods for electronic visual displays*. 9241 305, International Organization for Standardization, 2008.
- [ITU 94] ITU. *Specifications and alignment procedures for setting of brightness and contrast of displays*. Recommendation ITU-R BT.814-1, International Telecommunication Union, 1994.
- [ITU 98] ITU. *Subjective assessment methods for image quality in high-definition television*. Recommendation ITU-R BT.710-4, International Telecommunication Union, 1998.
- [ITU 04] ITU. *Methodology for the subjective assessment of the quality of television pictures*. Recommendation ITU-R BT.500-11, International Telecommunication Union, 2004.
- [ITU 05] ITU. *Report on results of comparative subjective picture quality assessment test between CRT and LCD*. Questions ITU-R 95/6, 102/6, International Telecommunication Union - Radiocommunication Study Groups, August 2005.
- [JVT 05] JVT. *H.264/Advanced Video Coding reference software version 10.1*. Joint Video Team, 2005. <http://iphone.hhi.de/suehring/tml/>.
- [Kelly 69] D. H. Kelly. *Flickering patterns and lateral inhibition*. Journal of the Optical Society of America, vol. 59, no. 10, pages 1361–1370, Octobre 1969.
- [Kelly 79a] D. H. Kelly. *Motion and vision. I. Stabilized images of stationary gratings*. Journal of the Optical Society of America A, vol. 69, no. 9, pages 1266–1274, September 1979.

- [Kelly 79b] D. H. Kelly. *Motion and vision. II. Stabilized spatio-temporal threshold surface*. Journal of the Optical Society of America A, vol. 69, no. 10, pages 1340–1349, October 1979.
- [Kelly 84] D. H. Kelly. *Retinal inhomogeneity. I. Spatiotemporal contrast sensitivity*. Journal of the Optical Society of America A, vol. 1, no. 1, pages 107–113, January 1984.
- [Kelly 85] D.H. Kelly. *Visual processing of moving stimuli*. Journal of the Optical Society of America A, vol. 2, no. 2, pages 216–225, February 1985.
- [Kim 07] T. Kim, B. Park, B.-H. Shin, H. Nam, B. H. Berkeley & S. S. Kim. *Analysis of response-time compensation for black-frame insertion*. Journal of the Society for Information Display, vol. 15, no. 9, pages 741–747, 2007.
- [Kim 09] Sang-Soo Kim, B. H. You, Cho J. H., D. G. Kim, B. H. Berkeley & N. D. Kim. *An 82-in. ultra-definition 120-Hz LCD TV using new driving scheme and advanced Super PVA technology*. Journal of the Society for Information Display, vol. 17, no. 2, pages 71–78, 2009.
- [Klompenshouwer 04] Michiel A. Klompenshouwer & Leo Jan Velthoven. *Motion blur reduction for liquid crystal displays : motion-compensated inverse filtering*. In Proceedings of the SPIE Conf. Visual Communications and Image Processing 2004, volume 5308, pages 690–699. Electronic Imaging 2006, January 2004.
- [Koenderink 78] Jan J. Koenderink, Maarten A. Bouman, Albert E. Bueno de Mesquita & Sybe Slappendel. *Perimetry of contrast detection thresholds of moving spatial sine wave patterns. I. The near peripheral visual field (eccentricity 0-8)*. Journal of the Optical Society of America, vol. 68, pages 845–849, 1978.
- [Koenderink 79] Jan J. Koenderink & Andrea J. van Doorn. *Spatiotemporal contrast detection threshold surface is bimodal*. Optics Letters, vol. 4, no. 1, pages 32–34, January 1979.
- [Kolb 02] Helga Kolb, Eduardo Fernandez & Ralph Nelson. *Webvision - The organization of the retina and visual system*. <http://webvision.med.utah.edu/>, 2002. University of Utah.
- [Kulikowski 76] J. J. Kulikowski. *Effective contrast constancy and linearity of contrast sensation*. Vision Research, vol. 16, no. 12, pages 1419–1431, December 1976.
- [Kumar 05] Sanjeev Kumar, Mainak Biswas & Truong Nguyen. *Analysis of the response time compensation for liquid crystal displays*. In Proceedings of European Signal Processing Conference, EUSIPCO, 2005.
- [Kurita 01] Taiichiro Kurita. *Moving picture quality improvement for hold-type AM-LCDs*. In SID Symposium Digest of Technical Papers, volume 32, pages 986–989. SID, June 2001.
- [Laird 06] Justin Laird, Mitchell Rosen, Jeff Pelz, Ethan Montag & Scott Daly. *Spatio-velocity CSF as a function of retinal velocity using unstabilized stimuli*. In Pro-

- ceedings of the SPIE Conf. Human Vision and Electronic Imaging XI, volume 6057. Electronic Imaging 2006, Janvier 2006.
- [LCDTVA 08] LCDTVA. *LCD TV matters, Volume 1, Issue 3*. LCD TV Association, Winter 2008.
- [Le Callet 01] Patrick Le Callet. *Critères objectifs avec référence de qualité visuelle des images couleur*. PhD thesis, Université de Nantes, 2001.
- [Ledgeway 95] Timothy Ledgeway & Andrew T. Smith. *The perceived speed of second-order motion and its dependence on stimulus contrast*. Vision Research, vol. 35, no. 10, pages 1421–1434, 1995.
- [Legge 87] Gordon E. Legge & D. Kersten. *Contrast discrimination in peripheral vision*. Journal of the Optical Society of America A, vol. 4, no. 8, pages 1594–1598, August 1987.
- [Lowry 61] E. M. Lowry & J. J. DePalma. *Sine-wave response of the visual system. I. The Mach phenomenon*. Journal of the Optical Society of America A, vol. 51, no. 7, pages 740–746, July 1961.
- [Manahilov 03] Velitchko Manahilov, Julie Calvert & William A. Simpson. *Temporal properties of the visual responses to luminance and contrast modulated noise*. Vision Research, vol. 43, pages 1855–1867, 2003.
- [McCartney 03] Richard I. McCartney. *A liquid crystal display response time compensation feature integrated into an LCD panel timing controller*. In SID Symposium Digest of Technical Papers, volume 34, pages 1350–1353. SID, May 2003.
- [Medjbeur 86] Soraya Medjbeur & Ulker Tulunay-Keesey. *Suprathreshold responses of the visual system in normals and in demyelinating diseases*. Investigative Ophthalmology and Visual Science, vol. 27, no. 9, pages 1368–1378, September 1986.
- [Monoyer 75] Ferdinand Monoyer. *Échelle typographique pour la détermination de l'acuité visuelle*. Comptes rendus 113, Académie des Sciences, 1875.
- [Munsell 15] Albert Henry Munsell. *Atlas of the munsell color system*. 1915.
- [Nachmias 67] Jacob Nachmias. *Effect of exposure duration on visual contrast sensitivity with square-wave gratings*. Journal of the Optical Society of America, vol. 57, no. 3, pages 421–427, March 1967.
- [Nipkow 84] Paul Nipkow. *Elektrisches Teleskop*. German Patent Office, 1884. Patent number : 30105.
- [Nose 01] T. Nose, M. Suzuki, D. Sasaki, M. Imai & H. Hayama. *A black stripe driving scheme for displaying motion pictures on LCDs*. In SID Symposium Digest of Technical Papers, volume 32, pages 994–997. SID, June 2001.
- [OCDE 00] OCDE. *Perspectives des technologies de l'information 2000 : Tic, commerce électronique et économie de l'information*. OECD Publishing, 2000.

- [Oka 04] Koichi Oka & Yoshi Enami. *Moving picture response time (MPRT) measurement system*. In SID Symposium Digest of Technical Papers, volume 35, pages 1266–1269. SID, May 2004.
- [Pan 05] Hao Pan, Xiao-Fan Feng & Scott Daly. *LCD motion blur modeling and analysis*. In IEEE International Conference on Image Processing, 2005. ICIP 2005., volume 2, pages 21–24, September 2005.
- [Péchar d 08] Stéphane Péchar d. *Qualité d’usage en télévision haute définition : évaluations subjectives et métriques objectives*. PhD thesis, Université de Nantes, 2008.
- [Reinitzer 88] Friedrich Reinitzer. *Beiträge zur Kenntniss des Cholesterins*. Monatshefte für Chemie, vol. 9, no. 1, pages 421–441, 1888.
- [Robson 66] J. G. Robson. *Spatial and temporal contrast-sensitivity functions of the visual system*. Journal of the Optical Society of America, vol. 56, no. Letters to the editor, pages 1141–1142, August 1966.
- [Schade 56] Otto H. Schade. *Optical and photoelectric analog of the eye*. Journal of the Optical Society of America, vol. 46, no. 9, pages 721–739, September 1956.
- [Shin 06] Kwang Hoon Shin, Ji Young Ahn, Ki Duk Kim, Hyun Ho Shin & In Jae Chung. *Acceptable motion blur levels of a LCD TV based on human visual system*. In SID Symposium Digest of Technical Papers, volume 37, pages 287–290. SID, June 2006.
- [Sluyterman 06] A. A. S. Sluyterman. *What is needed in LCD panels to achieve CRT-like motion portrayal ?* Journal of the Society for Information Display, vol. 14, no. 8, pages 681–, 2006.
- [Smith 73] Willoughby Smith. *Effect of Light on Selenium During the Passage of An Electric Current*. Nature, vol. 7, page 303, 1873.
- [Someya 07] Jun Someya & Hiroaki Sugiura. *Evaluation of liquid-crystal display motion blur with moving-picture response time and human perception*. Journal of the Society for Information Display, vol. 15, no. 1, pages 79–86, 2007.
- [Song 09] Wen Song, Kees Teunissen, Xiaohua Li, Yuning Zhang & Ingrid Heynderickx. *Evaluation of motion performance on scanning-backlight LCDs*. Journal of the Society for Information Display, vol. 17, no. 3, pages 251–261, 2009.
- [Spering 05] Miriam Spering, Dirk Kerzel, Doris I. Braun, Michael J. Hawken & Karl R. Gegenfurtner. *Effects of contrast on smooth pursuit eye movements*. Journal of Vision, vol. 5, pages 455–465, 2005.
- [Steinman 00] R. M. Steinman, Z. Pizlo & F. J. Pizlo. *Phi is not beta, and why Wertheimer’s discovery launched the Gestalt revolution : a minireview*. Vision Research, vol. 40, pages 2257–2264, 2000.
- [Stiles 59] W. S. Stiles & J. M. Burch. *NPL colour-matching investigation : Final report*. Optica Acta, vol. 6, pages 1–26, 1959.

- [Stockman 99] A. Stockman, L. T. Sharpe & C. C. Fach. *The spectral sensitivity of the human short-wavelength cones*. Vision Research, vol. 39, pages 2901–2927, 1999.
- [Stockman 00] A. Stockman & L. T. Sharpe. *Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype*. Vision Research, vol. 40, pages 1711–1737, 2000.
- [SVT 02] SVT. *Overall-quality assessment when targeting Wide XGA flat panel displays*. Rapport technique, SVT corporate development technology, 2002. [ftp ://ftp.ldv.e-technik.tu-muenchen.de/pub/test_sequences/](ftp://ftp.ldv.e-technik.tu-muenchen.de/pub/test_sequences/).
- [TCO 06] TCO. *TCO'06 Media Displays*. TCOF1076 version 1.2, TCO Development AB, Stockholm, Sweden, August 2006. http://www.allready.net/tcodevelopment1200/Datorer/TC006/TCOF1076_TC006_Media_Displays_1_2.pdf.
- [Teunissen 06] Kees Teunissen, Yuning Zhang, Xiaohua Li & Ingrid Heynderickx. *Method for predicting motion artifacts in matrix displays*. Journal of the Society for Information Display, vol. 14, no. 10, pages 957–964, 2006.
- [Trentacoste 07] Matthew Trentacoste, Wolfgang Heidrich, Lorne Whitehead, Helge Seetzen & Greg Ward. *Photometric image processing for high dynamic range displays*. Journal of Visual Communication and Image Representation, vol. 18, pages 439–451, 2007.
- [UKDisplay 09] UKDisplay. UK Displays and Lighting Newsletter, January 2009.
- [van Heesch 08] Franciscus H. van Heesch & Michiel A. Klompenhouwer. *The Temporal Aperture of Broadcast Video*. In SID Symposium Digest of Technical Papers, volume 1, pages 370–373, 2008.
- [van Santen 85] J. P. van Santen & G. Sperling. *Elaborated Reichardt detectors*. Journal of the Optical Society of America A, vol. 2, no. 2, pages 300–321, 1985.
- [VESA 05] VESA. *Flat Panel Display Measurements (FPDM2)*. Rapport technique 2.0 Update, Video Electronics Standards Association, May 2005.
- [Vikuiti 05] Vikuiti. *LCD Optics 101 (Multimedia Course)*. Vikuiti Internet site, Décembre 2005. http://solutions.3m.com/wps/portal/3M/en_US/Vikuiti1/BrandProducts/secondary/optics101/.
- [von Goethe 10] Johann W. von Goethe. *Traité des couleurs*. 1808-1810.
- [Wandell 95] Brian A. Wandell. *Foundations of vision*. Sinauer Associates, Sunderland MA, 1995.
- [Watanabe 68] A. Watanabe, T. Mori, S. Nagata & K. Hiwatashi. *Spatial sine-wave responses of the human visual system*. Vision Research, vol. 8, no. 9, pages 1245–1263, September 1968.
- [Watson 06] Andrew B. Watson. *The Spatial Standard Observer : A human vision model for display inspection*. In SID Symposium Digest of Technical Papers, volume 37, pages 1312–1315. SID, June 2006.

- [Waugh 94] Sarah J. Waugh & Robert F. Hess. *Suprathreshold temporal-frequency discrimination in the fovea and the periphery*. Journal of the Optical Society of America, vol. 11, pages 1199–1212, 1994.
- [Westheimer 72] Gerald Westheimer. Handbook of sensory physiology, volume 7, chapitre Visual acuity and spatial modulation thresholds, pages 170–187. Springer, 1972.
- [Wolf 99] Stephen Wolf & Margaret Pinson. *Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system*. In Proceedings of the SPIE Conference Multimedia Systems and Applications II, volume 3845, pages 266–277, 1999.
- [Wyszecki 82] G. Wyszecki & W. S. Stiles. Color science : concepts and methods, quantitative data and formulae (2nd ed.). New York : Wiley, 1982.
- [Yamamoto 06] Tsunenori Yamamoto, Susumu Sasaki, Yoshinori Tanaka & Youichi Igarashi. *Guiding principles for high-quality moving picture in LCD TVs*. Journal of the Society for Information Display, vol. 14, no. 10, pages 933–940, October 2006.
- [Yuyama 82] Ichiro Yuyama & Tetsuo Mitsuhashi. *Fundamental requirements for high-definition television systems*. NHK technical monograph, NHK, June 1982.
- [Zworykin 35] Vladimir K. Zworykin. *Method of and apparatus for producing images of objects*. United States Patent Office, November 1935. Patent number : 2021907.

Liste des publications

Revues spécialisées avec comité de lecture

Sylvain Tourancheau, Kjell Brunnström, Borje Andrén et Patrick Le Callet
LCD motion-blur measurements using different measurement methods
Journal of the Society for Information Display, Vol. 17, No. 3, pp. 239-249, mars 2009

Sylvain Tourancheau, Patrick Le Callet et Dominique Barba
Image and video quality assessment using LCD : comparisons with CRT conditions
IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences, Special Section on Image Media Quality, vol. E91-A, juin 2008

Conférences et manifestations internationales avec comité de lecture et publication des actes

Sylvain Tourancheau, Kjell Brunnström, Patrick Le Callet et Börje Andrén
Visual annoyance and user acceptance of LCD motion-blur
SID International Symposium, Display Week 2009, San Antonio (United States, TX), Juin 2009

Sylvain Tourancheau, Patrick Le Callet, Kjell Brunnström et Börje Andrén
Psychophysical study of LCD motion blur perception
Proceedings of the SPIE Conference Human Vision and Electronic Imaging XIV, HVEI2009, San Jose, USA, janvier 2009

Sylvain Tourancheau, Florent Autrusseau, Parvez Sazzad et Yuukou Horita
Impact of subjective dataset on the performance of image quality metrics
IEEE International Conference on Image Processing, ICIP2008, San Diego, USA, octobre 2008

Sylvain Tourancheau, Kjell Brunnström, Borje Andrén et Patrick Le Callet
Motion blur estimation on LCDs
Society for Information Display Symposium, SID2008, Los Angeles, USA, mai 2008

Sylvain Tourancheau, Patrick Le Callet et Dominique Barba

Impact of the resolution on the difference of perceptual video quality between CRT and LCD

IEEE International Conference on Image Processing, ICIP2007, San Antonio, USA, septembre 2007

Sylvain Tourancheau, Patrick Le Callet, Kjell Brunnström et Dominique Barba

Display awareness in subjective and objective video quality evaluation

European Signal Processing Conference, EUSIPCO 2007, Poznan, Pologne, septembre 2007

Patrick Le Callet, Stéphane Péchard, Sylvain Tourancheau, Alexandre Ninassi et Dominique Barba

Towards the next generation of video and image quality metrics : Impact of display, resolution, contents and visual attention in subjective assessment

Second International Workshop on Image Media Quality and its Applications, IMQA2007, Chiba, Japon, avril 2007

Sylvain Tourancheau, Patrick Le Callet et Dominique Barba

Influence of motion on contrast perception : supra-threshold spatio-velocity measurements

SPIE Conference Human Vision and Electronic Imaging XII, EI2007, San Jose, USA, janvier 2007

Sylvain Tourancheau, Patrick Le Callet et Dominique Barba

Prediction of perceived quality differences between CRT and LCD displays based on motion blur

Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM'07, Scottsdale, USA, janvier 2007

Stéphane Péchard, Sylvain Tourancheau, Patrick Le Callet, Mathieu Carnec et Dominique Barba

Towards video quality metrics for HDTV

Second International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM'06, Scottsdale, USA, janvier 2006

Colloque national avec publication des actes

Sylvain Tourancheau, Mathieu Carnec, Stéphane Péchard, Patrick Le Callet et Dominique Barba

Prédiction objective des différences de qualité perceptuelles entre un écran CRT et un écran LCD en TVHD

Compression et Représentation des Signaux Audiovisuels, CORESA2006, Caen, France, novembre 2006

Groupe de normalisation international

Sylvain Tourancheau, Stéphane Péchard, Patrick Le Callet, Mathieu Carnec, Arnaud Tirel et Dominique Barba

Motion blur : an explanation to subjective quality differences between CRT and LCD ?

Meeting of the Video Quality Experts Group (VQEG), Boston, USA, avril 2006

Groupe de recherche national

Sylvain Tourancheau, Stéphane Péchard, Mathieu Carnec, Patrick Le Callet et Dominique Barba
Évaluation de la qualité en télévision haute-définition

Journée du GDR Isis Thème B, Outils et méthodologies d'évaluation subjective de la qualité, Paris, France, juin 2006

Rapports de contrats

Dominique Barba, Mathieu Carnec, Patrick Le Callet, Stéphane Péchard, Arnaud Tirel et Sylvain Tourancheau

Human vision perception on HD displays : model

Projet européen ITEA/HD4U, Partie du livrable D3.4, décembre 2006

Dominique Barba, Mathieu Carnec, Patrick Le Callet, Stéphane Péchard, Arnaud Tirel et Sylvain Tourancheau

HDTV display-dependent elements for objective quality criterion

Projet européen ITEA/HD4U, Partie du livrable D2.4, septembre 2006

Dominique Barba, Mathieu Carnec, Patrick Le Callet, Stéphane Péchard, Arnaud Tirel et Sylvain Tourancheau

Human vision perception on LCD HDTV displays

Projet européen ITEA/HD4U, Partie du livrable D3.2, décembre 2005

Caractérisation objective et modélisation psychovisuelle du flou de mouvement sur les écrans à cristaux liquides. Impact sur la qualité perçue.

Résumé : La « haute-définition » a connu un essor important ces dernières années dans les domaines multimédia (télévision, jeux-vidéo, etc.). Cependant, pour que l'observateur bénéficie pleinement de cette amélioration de la qualité d'usage, l'augmentation de la résolution spatiale des images affichées doit être accompagnée d'une augmentation de la taille de l'écran. Les écrans à tube cathodique (CRT), lourds et encombrants, ont donc progressivement laissé la place à de nouvelles technologies d'affichage, caractérisées par leur écran plat. Les écrans à cristaux liquides (LCD), leaders sur le marché des téléviseurs et des moniteurs informatiques, souffrent encore de nombreux défauts visuels notamment lors de l'affichage d'images animées. Durant ces travaux de thèse, nous avons choisi de nous intéresser plus particulièrement au flou de mouvement, dont les causes sont à la fois technologiques et perceptives. Nous avons tout d'abord caractérisé cet artefact visuel d'un point de vue métrologique, en développant une méthode de mesure basée sur la réponse indicielle temporelle des pixels de l'écran. Dans un second temps, nous avons mis en œuvre plusieurs expérimentations psychophysiques de manière à modéliser la perception du flou de mouvement sur LCD, et à caractériser la gêne associée à ce défaut. Ces mesures psychovisuelles ont permis d'établir des recommandations sur les caractéristiques temporelles des écrans LCD. Enfin, un modèle de flou de mouvement global, basé sur la caractérisation technique et perceptive de cet artefact, permet de prédire la perte de qualité mesurée subjectivement sur les écrans LCD. Les solutions technologiques de réduction du flou de mouvement peuvent alors être évaluées de manière automatique.

Mots-clés : mesures de caractéristiques d'écran, perception visuelle, évaluation de la qualité visuelle, flou de mouvement objectif sur LCD, flou de mouvement perçu sur LCD, comparaison entre CRT et LCD.

Objective characterization and psychovisual modeling of liquid crystal displays motion-blur. Impact on the perceived quality.

Abstract : These last years, « high-definition » has been in rapid development in all multimedia domains (television, video gaming, etc.). However, in order to observers really enjoy this quality of experience improvement, the increase of the spatial resolution must come along with an increase of the display size. As a consequence, heavy and bulky cathode ray tube displays (CRT) have been progressively replaced by new flat-panel display technologies. Liquid crystal displays (LCD) are now leaders on the display market (TV sets and computer displays), but they are still suffering from visual artifacts, particularly when displaying moving pictures. In this PhD thesis, LCD motion-blur has been addressed. Its causes are both technological and perceptual. At first, this visual artifact has been characterized in metrologic way using measurement protocol based on the temporal step-response of the display pixels. Secondly, several psychophysics experimentations have been developed in order to model the LCD motion-blur perception and to characterize the annoyance experienced by the observers. From this psychovisual experiences, we have expressed some recommendations on the temporal characteristics of LC displays. Finally, a global motion-blur model, based on the perceptual characterization of this artifact, has been designed. It permits to predict the quality loss subjectively measured on LCD, and can be used to assess motion-blur reduction systems.

Keywords : display measurements, visual perception, visual quality assessment, objective LCD motion-blur, perceived LCD motion-blur, CRT and LCD comparison.
